

Ramsey properties of subsets of \mathbb{N}

Thesis submitted for the degree
“Doctor of Philosophy”

by

Alexander Fish

Submitted to the Senate of the Hebrew University

February 2, 2008

This work was carried out under the supervision of

Prof. Hillel Furstenberg

With my appreciation, I wish to thank my advisor Prof. Hillel Furstenberg for his guidance, support and encouragement during these unforgettable years of learning and research

Contents

Abstract	8
1 Introduction	10
1.1 Number theoretic aspects of Ramsey theory and Dynamics . . .	10
1.2 Generic points and WM sets	12
1.3 Examples of combinatorial properties of WM sets	14
1.4 Main results	16
1.4.1 Solvability of linear diophantine systems within WM sets	16
1.4.2 An additive analog of polynomial multiple recurrence for WM Sets	17
1.4.3 The equation $xy = z$ and normal sets	18
1.5 Structure of the thesis	19
2 Basic combinatorial properties of WM sets	20
2.1 Every WM set is a Poincaré set	20
2.2 Every WM set contains an IP set	21
2.3 Incomparability of Central and WM sets	22
3 Solvability of linear diophantine equations within WM sets	25
3.1 Proof of Sufficiency	25
3.2 Probabilistic constructions of WM sets	33
3.3 Comparison with Rado's Theorem	37
4 An additive analog of polynomial multiple recurrence for WM Sets . . .	39
4.1 Orthogonality of polynomial shifts	39
4.2 Proof of theorem 1.4.2	47
5 The equation $xy = z$ and normal sets	51
5.1 Normal sets and diophantine equations	51
5.2 A_Q is normal for a.e. Q	52
5.3 Solvability of the equation $xy = z^2$ and related problems	54
6 Appendix	56

Abstract

We associate ergodic properties to some subsets of the natural numbers. For any given family of subsets of the natural numbers one may study the question of occurrence of certain "algebraic patterns" in every subset in the family. By "algebraic pattern" we mean a set of solutions of a system of diophantine equations. In this work we investigate a concrete family of subsets - WM sets. These sets are characterized by the property that the dynamical systems associated to such sets are "weakly mixing", and as such they represent a broad family of randomly constructed subsets of \mathbb{N} . We find that certain systems of equations are solvable within every WM set, and our subject is to learn which systems have this property. We give a complete characterization of linear diophantine systems which are solvable within every WM set. In addition we study some non-linear equations and systems of equations with regard to the question of solvability within every WM set.

1 Introduction

The aim of the dissertation is to develop Ramsey theory as it relates to a special family of subsets of the natural numbers, namely, WM sets.

1.1 Number theoretic aspects of Ramsey theory and Dynamics

There are several domains in mathematics where the phenomena of Ramsey theory are encountered. The most classical one is graph theory. One of the best examples of Ramsey type theorems in graph theory is Ramsey's theorem:

For every $k \in \mathbb{N}$ there exists a natural number N big enough such that for every coloring into two colors of edges of the complete graph with N vertices there will exist a monochromatic complete subgraph with k vertices.

Throughout our work \mathbb{N} denotes the natural numbers.

Another theorem of the same spirit, where after a finite coloring of a structure we can find a substructure of the same type at least in one of colors, is van der Waerden theorem:

For every $r, l \in \mathbb{N}$ there exists $N(r, l) \in \mathbb{N}$ such that if the integers $\{1, 2, \dots, N(r, l)\}$ are partitioned into r sets, one of these contains arithmetic progressions of length $l + 1$.

Note that here $\{1, 2, \dots, N(r, l)\}$ may be replaced by any arithmetic progression of the same length.

Both Ramsey and van der Waerden theorems may be formulated in the following way:

After partitioning into a finite number of subsets of a "highly organized" structure (set) we will necessarily find one subset which contains the same substructure.

The difference between the two theorems is in the choice of "structure".

The foregoing finite version of van der Waerden theorem is equivalent to the following claim about finite partitions of the natural numbers:

For every partitioning of \mathbb{N} into a finite number of sets C_1, \dots, C_r at least one of the subsets contains arbitrarily long arithmetic progressions.

In the thirties of the twentieth century it was conjectured by Erdős and Turán that the pattern of arbitrarily long arithmetic progressions is not only stable for finite partitions but it necessarily appears in every subset of the natural numbers with positive upper Banach density. Later this conjecture was established by Szemerédi, see [16]:

The subsets of \mathbb{N} of positive upper Banach density contain arbitrarily long arithmetic progressions.

The structure of an arithmetic progression of length k can be viewed as a solution of the following diophantine system:

$$\begin{cases} x_2 - x_1 = x_3 - x_2 \\ x_2 - x_1 = x_4 - x_3 \\ \dots \\ x_2 - x_1 = x_k - x_{k-1}. \end{cases}$$

In this work we use extensively the notion of "algebraic pattern" or, to be more precise, we will speak of a subset S of natural numbers as containing some algebraic pattern. The latter means that for some diophantine system of equations in k variables, the set of solutions of the system intersects with S^k . Every pattern in this work will be an algebraic pattern. For example, an arithmetic progression of length k is an algebraic pattern.

There are algebraic patterns which are *regular* for finite partitions of \mathbb{N} ; i.e., one of the subsets of the partition necessarily contains the algebraic pattern, but no simple density condition implies that the pattern will be found. As an example of this we present Schur's theorem, [15]:

For every partitioning of \mathbb{N} into a finite number of sets C_1, \dots, C_r at least one of the subsets contains x, y, z such that $x+y=z$.

It is obvious that positivity of density for a subset S is not enough to ensure existence of a "Schur pattern" (e.g. $S = \text{odd numbers}$). In the context of van der Waerden and Schur theorems it will be appropriate to recall that there is a common generalization of them, Rado's theorem, which is a complete characterization of all linear patterns regular for finite partitions. By the word linear we mean that all equations in the diophantine system connected to the pattern are linear.

In the work we are motivated by the following question:

Are there conditions on $S \subset \mathbb{N}$ more restrictive than positive density that yield more algebraic patterns?

Our way to answer to the question is to add a condition of "random" behavior (which will be defined rigorously in the next subsection) to positivity of density of a subset. A subset which satisfies the foregoing two conditions (is called WM set) will contain Schur patterns. Here we would like to give a simple example of "random" behavior.

We recall that an infinite $\{0, 1\}$ -valued sequence λ is called a **normal sequence** if every finite binary word w occurs in λ with a right frequency $\frac{1}{2^{|w|}}$, where $|w|$ is a length of w . The more familiar notion is that of a normal number $x \in [0, 1]$. For every $x \in [0, 1]$, except a countable number of x 's, there exists a unique dyadic expansion: $x = \sum_{i=1}^{\infty} \frac{x_i}{2^i}$, $\forall i : x_i \in \{0, 1\}$. Then x is called a normal number if and only if the sequence $(x_1, x_2, \dots, x_n, \dots)$ is a normal sequence. To a sequence $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n, \dots) \in \{0, 1\}^{\mathbb{N}}$ we associate the set $B_\lambda \subset \mathbb{N}$ by the rule: $i \in B_\lambda \leftrightarrow \lambda_i = 1$. We define the notion of a normal set.

A set $S \subset \mathbb{N}$ is called **normal** if there exists a normal sequence $\lambda \in \{0, 1\}^{\mathbb{N}}$ such that $B_\lambda = S$.

Normal sets exhibit a non-periodic, "random" behavior. We remark that every normal set contains Schur patterns. We notice that if S is a normal set then $S - S$ contains \mathbb{N} . Therefore, the equation $z - y = x$ is solvable within every normal set. From the last statement it follows that every normal set contains Schur patterns.

We are looking for a possible answer to the foregoing question by using a dynamical approach. All aforementioned theorems have dynamical equivalent formulations. For our question the most relevant theorem is Szemerédi's theorem. Furstenberg has shown that Szemerédi's theorem is equivalent to the phenomenon of multiple recurrence valid for general volume preserving dynamical systems which can be established by purely dynamical techniques (see [10]).

In this context Furstenberg formulates a correspondence principle for subsets of the natural numbers of positive upper Banach density:

Given a set $E \subset \mathbb{N}$ with $d^(E) > 0$ (E of positive upper Banach density) there exists a probability measure preserving system (X, \mathbb{B}, μ, T) and a set $A \in \mathbb{B}$, $\mu(A) = d^*(E)$, such that for any $k \in \mathbb{N}$ and any $n_1, \dots, n_k \in \mathbb{Z}$ one has:*

$$d^*(E \cap (E - n_1) \cap \dots \cap (E - n_k)) \geq \mu(A \cap T^{-n_1}A \cap \dots \cap T^{-n_k}A).$$

By this correspondence principle in order to prove Szemerédi's theorem it is sufficient to establish the following multiple recurrence theorem which is proved purely dynamically in [10].

For any probability measure preserving system (X, \mathbb{B}, μ, T) , a set $A \in \mathbb{B}$, $\mu(A) > 0$ and any $k \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that $\mu(A \cap T^{-n}A \cap T^{-2n}A \cap \dots \cap T^{-(k-1)n}A) > 0$.

The basic idea of the correspondence principle is that a set of positive density can be viewed more or less (there are some technicalities) as return times of generic points of ergodic systems to a set of positive measure. If a dynamical system will be even more "random" (for example weakly mixing or mixing) then we expect to find that within a set of return times to a set of positive measure one can find a greater variety of algebraic patterns.

Our approach is to deal with the sets of integers that are the return times of a generic point of weakly mixing system to a set of positive measure. Such subsets of \mathbb{N} we call WM sets. We formalize this in the next section.

1.2 Generic points and WM sets

To define formally the main object of this work we need the notions of measure preserving systems and of generic points.

Definition 1.2.1 *Let X be a compact metric space, \mathbb{B} be the Borel σ -algebra on X , let $T : X \rightarrow X$ be a measurable map and μ a probability measure on \mathbb{B} . A quadruple (X, \mathbb{B}, μ, T) is called a **measure preserving system** if for every $B \in \mathbb{B}$ we have $\mu(T^{-1}B) = \mu(B)$.*

For a compact metric space X we denote by $C(X)$ the space of continuous functions on X with the uniform norm.

Definition 1.2.2 Let (X, \mathbb{B}, μ, T) be a measure preserving system. A point $\xi \in X$ is called **generic** if for any $f \in C(X)$ we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n \xi) = \int_X f(x) d\mu(x). \quad (1.1)$$

We can now give an alternative definition of a normal set which is purely dynamical. A set S is normal if and only if the sequence $1_S \in \{0, 1\}^{\mathbb{N}}$ is a generic point of the measure preserving system $(\{0, 1\}^{\mathbb{N}}, \mathbb{B}, T, \mu)$, where \mathbb{B} is Borel σ -algebra on the topological space $\{0, 1\}^{\mathbb{N}}$ which is endowed with the Tychonoff topology, T is the shift to the left, μ is the product measure of μ_i 's where $\mu_i(0) = \mu_i(1) = \frac{1}{2}$. Thus, the system $(\{0, 1\}^{\mathbb{N}}, \mathbb{B}, T, \mu)$ is the Bernoulli $(\frac{1}{2}, \frac{1}{2})$ system and, in particular, it is a mixing dynamical system.

The notion of a WM set generalizes that of a normal set, where the role played by a Bernoulli dynamical system is taken over by dynamical systems of more general character.

Let $\xi(n)$ be any $\{0, 1\}$ -valued sequence. There is a natural dynamical system (X_ξ, T) connected to the sequence ξ :

On the foregoing compact space $\Omega = \{0, 1\}^{\mathbb{N}}$ which is endowed with the Tychonoff topology, we define a continuous map $T : \Omega \rightarrow \Omega$ by shifting all the elements of a sequence to left, namely, $(T\omega)_n = \omega_{n+1}$. Now for any ξ in Ω we define X_ξ to be $\overline{(T^n \xi)_{n \in \mathbb{N}}} \subset \Omega$.

Let A be a subset of \mathbb{N} . Choose $\xi = 1_A$ and assume that for an appropriate measure μ , the point ξ is generic for $(X_\xi, \mathbb{B}, \mu, T)$. Now we attach to the set A dynamical properties associated with the system $(X_\xi, \mathbb{B}, \mu, T)$.

For example, A is called weakly mixing (respectively - totally ergodic) if the measure preserving system $(X_\xi, \mathbb{B}, \mu, T)$ is weakly mixing (respectively - totally ergodic).

We recall the latter two notions of ergodic theory.

Definition 1.2.3 A measure preserving system (X, \mathbb{B}, μ, T) is called **ergodic** if every $A \in \mathbb{B}$ which is invariant under T , i.e. $T^{-1}(A) = A$, satisfies $\mu(A) = 0$ or 1 .

A measure preserving system (X, \mathbb{B}, μ, T) is called **totally ergodic** if for every $n \in \mathbb{N}$ the system $(X, \mathbb{B}, \mu, T^n)$ is ergodic.

A measure preserving system (X, \mathbb{B}, μ, T) is called **weakly mixing** if the system $(X \times X, \mathbb{B}_{X \times X}, \mu \times \mu, T \times T)$ is ergodic.

Let \mathcal{P} denote some dynamical property of a measure preserving system. We can attach the property \mathcal{P} to a subset of the natural numbers by the following:

Definition 1.2.4 A subset $S \subset \mathbb{N}$ is \mathcal{P} $\Leftrightarrow 1_S$ is generic for measure preserving system $(X_{1_S}, \mathbb{B}, \mu, T)$ which has property \mathcal{P} .

Finally, we would like to deal with subsets of \mathbb{N} which may have a rich structure, i.e. may be expected to exhibit many algebraic patterns. Therefore, we restrict ourselves to the case of weakly mixing subsets of \mathbb{N} of positive density (the density of every weakly mixing set exists!). For completeness we define the density of a subset of \mathbb{N} .

Definition 1.2.5 Let $S \subset \mathbb{N}$. If the limit of $\frac{1}{N} \sum_{n=1}^N 1_S(n)$ exists as $N \rightarrow \infty$ we call it the **density** of S and denote by $d(S)$.

Definition 1.2.6 A subset $S \subset \mathbb{N}$ is called a **WM set** if S is weakly mixing and the density of S is positive. That is to say, 1_S is a generic point of the weakly mixing system $(X_{1_S}, \mathbb{B}, \mu, T)$ and $d(S) > 0$.

We could equally well speak of strongly mixing sets, but for our purposes, weak mixing will be adequate.

1.3 Examples of combinatorial properties of WM sets

We would like to list basic combinatorial/Ramsey properties of WM sets. To do this we recall the definitions of two basic notions in ergodic Ramsey theory and combinatorial number theory.

Definition 1.3.1 A set $S \subset \mathbb{N}$ is called a **Poincaré set** if for every measure preserving system (X, Σ, T, μ) (not necessarily topological system) and every $A \in \Sigma$ with $\mu(A) > 0$ there exists $n \in S$ such that $\mu(A \cap T^{-n}A) > 0$.

This can be reformulated in purely combinatorial terms. First, we recall the notion of upper Banach density for a subset of the natural numbers.

Definition 1.3.2 Let $E \subset \mathbb{N}$. Upper Banach density of E , $d^*(E)$ is the following quantity

$$d^*(E) = \limsup_{b_n - a_n \rightarrow \infty} \frac{|E \cap \{a_n, \dots, b_n\}|}{b_n - a_n + 1}.$$

By Furstenberg's correspondence principle, for a set S to be Poincaré is equivalent to the following:

For every subset E of positive upper Banach density there exists $s \in S$ with $d^*(E \cap (E - s)) > 0$.

In fact, a milder condition is sufficient: a set S is Poincaré if and only if for every E of positive upper Banach density there exists $s \in S$ such that $E \cap (E - s) \neq \emptyset$; that is to say that s is a difference of two numbers in E .

The last property is called 1-recurrence. Sometimes in the literature a Poincaré set is called 1-recurrent set.

The next notion is taken from combinatorial number theory and may be viewed as a generalization of an infinite arithmetic progression around 0.

Definition 1.3.3 A set $S \subset \mathbb{N}$ is called **IP-set** if there exists an infinite sequence of natural numbers $\{p_1, p_2, \dots, p_n, \dots\}$ (not necessarily different) such that

$$S = \{p_{i_1} + \dots + p_{i_k} \mid i_1 < i_2 < \dots < i_k, k \in \mathbb{N}\}.$$

We recall the definition of IP*-set.

Definition 1.3.4 *A set S is called an IP*-set if for every IP-set E we have $E \cap S \neq \emptyset$.*

The next two results which will be proved in Chapter 2 give a first evidence of the richness of algebraic patterns which occur in every WM set.

Theorem 1.3.1 *Every WM set is a 1-recurrent set. (\Rightarrow Poincaré set)*

Theorem 1.3.2 *Every WM set contains an IP-set.*

Corollary 1.3.1 *A non-trivial WM set (which has density less than 1) is never an IP*-set.*

One of the reasons to choose WM sets as an object of our research and not sets which satisfy weaker conditions, for example, totally ergodic sets, is the fact that the foregoing theorems don't hold for totally ergodic sets. The following is an example of a totally ergodic set which is neither a Poincaré set nor contains an IP-set.

Example 1 *Let $\alpha \notin \mathbb{Q}$ and denote by S the following subset of \mathbb{N}*

$$S = \left\{ n \in \mathbb{N} \mid \alpha n \bmod 1 \in \left[\frac{2}{5}, \frac{3}{5} \right] \right\}.$$

Then S is a totally ergodic set of positive density which is not Poincaré set and for every $x, y \in S$ we have $x + y \notin S$.

Proof. We start from the last statement which is easily proven. Namely, if $\xi, \eta \in [\frac{2}{5}, \frac{3}{5}]$ then $(\xi + \eta) \bmod 1 \notin [\frac{2}{5}, \frac{3}{5}]$. It follows that if $x, y \in S$ then $(x + y)\alpha \bmod 1 \notin [\frac{2}{5}, \frac{3}{5}]$ and therefore $x + y \notin S$. This implies S contains no IP-set.

S is not a Poincaré set, as we see by checking the recurrence condition of definition 1.3.1 for the system $(\mathbb{T}, \mathbb{B}, S_\alpha, \lambda)$, where \mathbb{T} is the one dimensional torus, $S_\alpha(x) = x + \alpha$, λ is lebesgue measure and the subset $A = [0, \frac{1}{5}]$ is of measure $\frac{1}{5}$. Then obviously for every $s \in S$ we have $\lambda(A \cap S_\alpha^{-s}A) = 0$.

To show that S is a totally ergodic set we note that S consists of return times to the set $I = [\frac{2}{5}, \frac{3}{5}]$ of the point zero within the aforementioned measure preserving system $(\mathbb{T}, \mathbb{B}, S_\alpha, \lambda)$. Consider the space of $\{0, 1\}$ -sequences X_{1_S} , and consider the characteristic function $\chi \in C(X_{1_S})$ of a cylinder

$$C_{i_1, \dots, i_k}^{j_1, \dots, j_k} = \{\omega \in \{0, 1\}^\infty \mid \omega_{i_l} = j_l, \forall 1 \leq l \leq k\}.$$

We have

$$\frac{1}{N} \sum_{n=1}^N \chi(T^n 1_S) = \frac{1}{N} \sum_{n=1}^N \phi_{j_1}(n + i_1) \dots \phi_{j_k}(n + i_k) =$$

$$\frac{1}{N} \sum_{n=1}^N T^n f(0) \rightarrow_{N \rightarrow \infty} \int_{\mathbb{T}} f(x) d\lambda(x),$$

where $\phi_1(n) = 1_I(\alpha n)$, $\phi_0(n) = 1 - 1_I(\alpha n)$ and $f(x) = \prod_{i=1}^k T^{i_i} \phi_{j_i}(x)$. Since the linear space of characteristic functions on cylinders is dense in $C(X_{1_S})$ we conclude that the point 1_S is a generic point in X_{1_S} for a measure which is obtained as a projection of lebesgue measure in $(\mathbb{T}, \mathbb{B}, S_\alpha, \lambda)$. Thus our system is a factor of a totally ergodic system $(\mathbb{T}, \mathbb{B}, S_\alpha, \lambda)$; therefore it is itself totally ergodic.

□

The concept of WM sets is new. It relies on properties of the corresponding point 1_S within a measure preserving dynamical system. There is a concept of "good" subsets in the context topological dynamics (they contain many algebraic patterns) which is defined by H. Furstenberg (see [11]); namely, *central sets*. In order to define these sets we define a uniformly recurrent point in a topological dynamical system.

Definition 1.3.5 *Let (X, T) be a topological dynamical system, i.e., X is a metric compact space and $T : X \rightarrow X$ is a continuous transformation. A point $x_0 \in X$ is called **uniformly recurrent** if for any open set U , such that $x_0 \in U$, the set $\{n \in \mathbb{N} | T^n x_0 \in U\}$ is syndetic (a set with bounded gaps).*

Definition 1.3.6 *Let (X, T) be a topological dynamical system. Denote by d a metric on X . Two points $x, y \in X$ will be called **proximal** if there exists an increasing sequence $\{n_k\}$ such that $\lim_{k \rightarrow \infty} d(T^{n_k} x, T^{n_k} y) = 0$.*

Definition 1.3.7 *A set $S \subset \mathbb{N}$ is called **central** if there exists a topological dynamical system (X, T) , a uniformly recurrent point $x_0 \in X$, a point $x \in X$ which is proximal to x_0 and a neighborhood U of x_0 such that $S = \{n \in \mathbb{N} | T^n x \in U\}$.*

In the section 2.3 we prove the incomparability of central and WM sets.

Theorem 1.3.3 *There exists a WM set which does not contain a central set.*

Remark 1.3.1 The opposite direction is easy; for example, we could take the set of even numbers.

1.4 Main results

1.4.1 Solvability of linear diophantine systems within WM sets

We have succeeded to give a complete characterization of those linear systems of diophantine equations which are solvable within every WM set.

Theorem 1.4.1 *Let $B \in \mathbb{Q}^{t \times k}$ and $\vec{d} \in \mathbb{Q}^t$. The system of linear equations*

$$B\vec{x} = \vec{d} \tag{1.2}$$

is solvable within every WM set \Leftrightarrow there exist three vectors $\vec{x}_1 = (a_1, a_2, \dots, a_k)^t$, $\vec{x}_2 = (b_1, b_2, \dots, b_k)^t$, $\vec{f} = \{f_1, f_2, \dots, f_k\}^t \in \mathbb{N}^k$, disjoint sets $E, F_1, \dots, F_l \subset \{1, 2, \dots, k\}$,

$E \cup F_1 \cup \dots \cup F_l = \{1, 2, \dots, k\}$, such that:

a) for every $i, j \in E$, $i \neq j$

$$\det \begin{pmatrix} a_i & b_i \\ a_j & b_j \end{pmatrix} \neq 0.$$

b) for every $p \in \{1, \dots, l\}$ there exist $c_1^p, c_2^p \in \mathbb{N}$, such that for every $i \in F_p$ we have $a_i = c_1^p$, $b_i = c_2^p$ and for every $j \in \{1, \dots, k\} \setminus F_p$ we require

$$\det \begin{pmatrix} a_j & b_j \\ c_1^p & c_2^p \end{pmatrix} \neq 0.$$

c) The vector \vec{f} is constant on all indices from the same F_p with $p \in \{1, 2, \dots, l\}$, namely,

$$\forall p \in \{1, 2, \dots, l\} \forall i \in F_p : f_i = f^p,$$

where $f^p \in \mathbb{Z}$. d) The affine space of solutions of the system $B\vec{x} = \vec{d}$ contains

$$\{n\vec{x}_1 + m\vec{x}_2 + \vec{f} \mid n, m \in \mathbb{N}\}.$$

A proof of this theorem is in section 3 of the thesis.

As we will show in proposition 3.3.1, it will follow from theorem 1.4.1 that every linear algebraic pattern which is regular for finite partitions (Rado theorem gives a complete characterization of such patterns) occurs in every WM set. In the context of Rado patterns and WM sets it is important to recall that by Furstenberg's theorem (see [11]) every central set contains all Rado patterns. By theorem 1.3.3 it follows that we can't prove that every WM set contains all Rado patterns by use of Furstenberg's theorem.

1.4.2 An additive analog of polynomial multiple recurrence for WM Sets

A natural generalization of the theorem of Szemerédi is the seminal theorem of Bergelson and Leibman about polynomial multiple recurrence [5]. If we rephrase this theorem combinatorially it states that for every k polynomials which are essentially distinct (i.e., no two differ by a constant) p_1, \dots, p_k with positive leading coefficients and $p_1(0) = p_2(0) = \dots = p_k(0) = 0$, and for every subset A of the natural numbers of positive upper Banach density, there exists $n \in \mathbb{N}$ such that $\{x, x + p_1(n), \dots, x + p_k(n)\} \in A^{k+1}$. The latter means that the system of equations

$$\begin{cases} y_1 - x = p_1(n) \\ \dots \\ y_k - x = p_k(n) \end{cases}$$

is solvable in every set of positive upper Banach density for some $n \in \mathbb{N}$. For A a WM set we can use Bergelson's PET theorem (see [2]) and to obtain the same result without the restriction that all polynomials have zero free coefficient. If additionally we require that $n \in A$ then we can use the IP-polynomial Szemerédi theorem of Bergelson, Furstenberg and McCutcheon (see [3]) and the fact that every WM set contains an IP-set to establish that the previous system is solvable within every WM set A and $n \in A$

provided $p_i(0) = 0, \forall i : 1 \leq i \leq k$. It is very natural question to try to establish the analogous result for the "additive" system which is obtained from the last one by replacing all minuses by pluses.

$$\begin{cases} x + y_1 = p_1(z) \\ x + y_2 = p_2(z) \\ \dots \\ x + y_k = p_k(z) \end{cases} \quad (1.3)$$

Of course, in the case of the additive system we can't expect that there exists a solution within every set of positive upper density (there are a lot of examples of periodic sets that contain no solution for the equation $x + y = n^2$; i.e., the set $5\mathbb{N} + 1$). On the other hand, we would expect that for some such systems there exists a solution within every WM set, where congruence conditions do not form an obstruction. We can obtain the following characterization of solvability of system (1.3) within every WM set.

Theorem 1.4.2 *For every $k \in \mathbb{N}$ the system (1.3) is solvable within every WM set if $\deg(p_1) = \deg(p_2) = \dots = \deg(p_k)$, the difference of every two polynomials is a non-constant polynomial and all leading coefficients of p_1, \dots, p_k are positive.*

There is an easy case which shows the necessity of some restrictions on the degrees of the polynomials; namely, when in the system (1.3) there are two polynomials with degrees which differ by at least two.

Remark 1.4.1 If in the system (1.3) there are two polynomials with degrees which differ by at least two, then there exist WM sets within which the system (1.3) is unsolvable.

Proof. We take an arbitrary WM set A ; then removing a set of density zero from A leads again to a WM set. In particular, we can exclude from A all solutions of the system (1.3) by removing a set of density zero. Namely, if $\deg p_1 \leq \deg p_2 - 2$ then replace A by

$$A' = A \setminus \left(\bigcup_{n \in \mathbb{N}} [p_2(n) - p_1(n), p_2(n)] \right)$$

which is again a WM set. (For sufficiently large n the polynomials $p_1(n), p_2(n)$ are monotone.) Within A' the system (1.3) is unsolvable.

□

1.4.3 The equation $xy = z$ and normal sets

We recall the notion of a normal set.

We have the natural bijection between infinite binary $\{0, 1\}$ -sequences and subsets of \mathbb{N} , namely for any sequence λ we associate the subset $B_\lambda = \{i | \lambda_i = 1\}$.

Definition 1.4.1 *A set $B \subset \mathbb{N}$ is called **normal** if the infinite binary sequence λ which corresponds to B (i.e. $B_\lambda = B$) is normal.*

In the section 5 we prove the following result.

Theorem 1.4.3 *There exist normal sets within which the equation $xy = z$ is unsolvable.*

Our proof is non-constructive and we do not know an explicit example.

On the other hand the equation $xy = z^2$ is solvable in any normal set, and in fact:

Theorem 1.4.4 *Let $A \subset \mathbb{N}$ be a WM set. Then there exist $x, y, z \in A$ ($x \neq y$) such that $xy = z^2$.*

For normal sets we can also show the following

Theorem 1.4.5 *Let $A \subset \mathbb{N}$ be an arbitrary normal set. Then there exist $x, y, u, v \in A$ such that $x^2 + y^2 = \text{square}$ and $u^2 - v^2 = \text{square}$.*

This result holds for WM sets as well provided their density exceeds $\frac{1}{3}$.

1.5 Structure of the thesis

The thesis consists of 5 sections and an Appendix. The first section is an introduction to the subject of the thesis, namely WM sets, and a formulation of main results. In the second section we prove basic combinatorial properties of WM sets, which rely on 1-recurrence of WM sets. In addition we show that the notions of central sets and of WM sets are incomparable. In the third section we give a proof of the theorem which characterizes all linear diophantine systems which are solvable within every WM set. In the fourth section we prove that the system (1.3) is solvable within every WM set if all the polynomials are essentially distinct, have the same degree and have positive leading coefficients. The section 5 is devoted to non-linear equations. In particular, we show the existence of a normal set for which the (non-linear) equation $xy = z$ has no solutions with x, y, z in the set. In Appendix we collected some technical lemmas which are used in more than one section. In particular we formulate and prove the van der Corput lemma which will be used on several occasions.

2 Basic combinatorial properties of WM sets

2.1 Every WM set is a Poincaré set

In this section we prove the following theorem.

Theorem 1.3.1 *Every WM set is a 1-recurrent set. (\Rightarrow Poincaré set)*

To prove theorem 1.3.1 we note that it is sufficient by the ergodic decomposition theorem to show recurrence of a WM set for every ergodic system. We show the following

Proposition 2.1.1 *Let S be a WM set. Then for every ergodic measure preserving system (X, Σ, μ, T) and every $A \in \Sigma$ with $\mu(A) > 0$ there exists $s \in S$ such that $\mu(A \cap T^{-s}A) > 0$.*

Proof. We make use of spectral theory. Namely, by spectral theory for the unitary operator $U : L^2(X, \Sigma, \mu) \rightarrow L^2(X, \Sigma, \mu)$ which is defined by $Uf = f \circ T$ and the function $1_A \in L^2(X, \mu)$ there exists a spectral measure ω_{1_A} (we denote it simply ω) on \mathbb{T} (the spectrum of U) such that for every $n \in \mathbb{N}$ we have

$$\langle 1_A, T^n 1_A \rangle = \int_{[0,1]} e^{2\pi i x n} d\omega(x).$$

Let \mathbb{T} denotes 1-dimensional torus and for every $\alpha : 0 \leq \alpha < 1$ let

$S_\alpha(x) \doteq x + \alpha \pmod{1}$. For every $\alpha \in (0, 1)$ consider the Kronecker system (K, S_α) , where $K = \overline{\{S_\alpha^n(0)\}_{n=0}^\infty}$. For $\alpha \notin \mathbb{Q}$ this system is a factor of the system $(\mathbb{T}, \mathbb{B}, S_\alpha, \lambda)$, defined in § 1.3, and in any case the Kronecker system is disjoint from the weak-mixing system $(\overline{\{T^n 1_S\}_{n=1}^\infty}, \mathbb{B}, \mu, T)$. We use the theorem of Furstenberg:

If the measure preserving systems $(X, \mathbb{B}_X, \mu, T_X)$ and $(Y, \mathbb{B}_Y, \nu, T_Y)$ are disjoint, $x \in X$ and $y \in Y$ are generic then $(x, y) \in X \times Y$ is generic for the system $(X \times Y, \mathbb{B}_X \times \mathbb{B}_Y, \mu \times \nu, T_X \times T_Y)$ (see [9]).

Applying this to the pair $(0, 1_S) \in \mathbb{T} \times \overline{\{T^n 1_S\}}$ we obtain

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N 1_S(n) e^{2\pi i \alpha n} = 0. \quad (2.1)$$

Therefore, by use of Lebesgue dominated convergence theorem from (2.1) we have

$$\frac{1}{N} \sum_{n=1}^N 1_S(n) \langle 1_A, T^n 1_A \rangle \rightarrow_{N \rightarrow \infty} d(S)\omega(0).$$

But $\omega(0)$ is represented in terms of integral over 1_A by the following

$$\begin{aligned} \omega(0) &= \int_{[0,1]} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i x n} d\omega(x) = \\ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \int_{[0,1]} e^{2\pi i x n} d\omega(x) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \langle 1_A, T^n 1_A \rangle = \langle 1_A, \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N T^n 1_A \rangle = \end{aligned}$$

$$\left(\int_X 1_A(x) d\mu(x) \right)^2 = \mu(A)^2.$$

We have used ergodicity of the system X in the last step.

Finally, we get

$$\frac{1}{N} \sum_{n=1}^N 1_S(n) < 1_A, T^n 1_A > \xrightarrow{N \rightarrow \infty} d(S) \mu(A)^2 > 0.$$

Since the inner product $< 1_A, T^n 1_A > = \mu(A \cap T^{-n} A)$, we conclude that there exists $s \in S$, such that $\mu(A \cap T^{-s} A) > 0$.

□

2.2 Every WM set contains an IP set

To prove that every WM set contains an IP set we use theorem 1.3.1.

Proof. (of theorem 1.3.2)

Let S be an arbitrary WM set. By use of theorem 1.3.1 we conclude that there exists $s \in S$ such that $(S - s) \cap S$ has positive density. We shall see that it is again a WM set.

To prove the last statement we define in the weak-mixing measure preserving space $(X = \overline{\{T^n 1_S\}_{n=1}^\infty}, \mathbb{B}, \mu, T)$ the set $A = \{x \in X \mid (x)_0 = 1\}$. Then $\mu(A) = d(S) > 0$ (by use of genericity of 1_S in X) and by using recurrence of the set S we obtain that there exists $s \in S$, such that $\mu(A \cap T^{-s} A) > 0$. By genericity of the point $1_S \in X$ it follows that $\mu(A \cap T^{-s} A) = d(S \cap (S - s))$. The map $\phi : X \rightarrow \{0, 1\}^\mathbb{N}$ defined by $\phi(x) = y$, where $y(n) = x(n)x(n + s)$ takes X to a closed shift invariant set Y in $\{0, 1\}^\mathbb{N}$ and $\phi(1_S) = 1_{S \cap (S - s)}$.

Therefore the point $1_{S \cap (S - s)} \in \{0, 1\}^\mathbb{N}$ is a generic point of (Y, T) which is again a weak-mixing measure preserving system. Here we get weak-mixing because the resulting system is a factor of the system (X, \mathbb{B}, μ, T) .

The next stage of our proof is to define inductively an IP set in S .

Let $s_1 \in S$, such that $S \cap (S - s_1)$ is again a WM set.

If we denote by $S_1 = S \cap (S - s_1)$ (a WM set) then we define $s_2 \in S_1$, such that $S_1 \cap (S_1 - s_2)$ is again a WM set. Note that if $s_3 \in S_1 \cap (S_1 - s_2)$ then $s_3 + s_1, s_3 + s_2, s_3 + s_1 + s_2 \in S$.

If we have defined s_1, \dots, s_n and a WM set S_n we define the element $s_{n+1} \in S_n$ and a WM set S_{n+1} by the following: there exists an element $s_{n+1} \in S_n$, such that $S_n \cap (S_n - s_{n+1})$ is again a WM set, which we denote by S_{n+1} .

In this way we have defined an infinite sequence $\{s_1, s_2, \dots, s_n, \dots\} \subset S$. It is a consequence of the construction of the sequence that every finite sum of its elements is again in S .

Therefore we have found an IP set within an arbitrary WM set.

□

2.3 Incomparability of Central and WM sets

We will use a variant of Rohlin's lemma in ergodic theory.

Lemma 2.3.1 *Let (X, \mathbb{B}, μ, T) be an ergodic non-periodic invertible measure preserving system (m.p.s.). Then for any $n > 1$ there exists $C \in \mathbb{B}$ with $\frac{1}{n} \leq \mu(C) \leq \frac{1}{n-1}$, such that $X = \bigcup_{i=0}^n T^i C$, where the equality is up to a measure zero set and every point $x \in \bigcup_{i=0}^n T^i C$ returns to C by at most $n + 1$ iterations of T .*

Proof. Let us fix $n > 1$. For any $B \in \mathbb{B}$ with $\mu(B) > 0$ let us build up Kakutani's tower, by the following procedure.

Denote by $B_k = \{x \in B \mid r_B(x) = k\}$, where $r_B(x) = \min_{i \geq 1} \{i \mid T^i x \in B\}$. By the Poincaré recurrence theorem we have $B = \bigcup_{i=1}^{\infty} B_i$. Then the following family of sets $B_1, (B_2 \cup TB_2), \dots, (B_k \cup \dots \cup T^{k-1}B_k), \dots$ is called Kakutani's tower with base B . Obviously by ergodicity it follows $X = \bigcup_{k=1}^{\infty} \bigcup_{i=0}^{k-1} T^i B_k$ (where the equality is up to a set of measure zero) and the union is measurably disjoint (this means that the intersection of any two sets from the union is of measure zero).

We introduce $C = \bigcup_{m=0}^{\infty} \bigcup_{k=m(n+1)+1}^{\infty} T^{m(n+1)} B_k$. Then $X = \bigcup_{i=0}^n T^i C$, from which we get the estimation $\frac{1}{n} \leq \mu(C)$, and any point in C returns back to C by at most $n + 1$ iterations of T . On the other hand if we denote by B' the higher layer of C , then obviously we have

$$\mu(C \setminus B') \leq \frac{1}{n}.$$

In addition we have $\mu(B') = \mu(B)$. Therefore we get

$$\mu(C) \leq \frac{1}{n} + \mu(B).$$

Now let us choose $B \in \mathbb{B}$ such that $0 < \mu(B) < \frac{1}{n-1} - \frac{1}{n} = \frac{1}{n(n-1)}$ (it can be done because our m.p.s. is non-periodic and is non-atomic). Finally with appropriate choice of B we get the desired result.

□

Remark 2.3.1 *We don't have to assume that the system is non-periodic; it is sufficient that for any $\varepsilon > 0$ there exists $B \in \mathbb{B}$ with $0 < \mu(B) \leq \varepsilon$.*

Remark 2.3.2 *Every weak-mixing m.p.s. is non-periodic.*

Definition 2.3.1 *A sequence $\omega \in \{0, 1\}^{\mathbb{N}}$ is said to satisfy **property** (l, L) if for any $l > 0$ there exists $L > 0$ such that for every $k \geq 0$ the block $(\omega_k, \dots, \omega_{k+L-1})$ contains at least one subblock of l successive zeros.*

We attach the notion of centrality to $\{0, 1\}$ -valued sequences as well.

Definition 2.3.2 *A sequence $\lambda \in \{0, 1\}^{\mathbb{N}}$ is called **central** if the set $B_\lambda = \{i \in \mathbb{N} \mid \lambda_i = 1\}$ is central.*

Lemma 2.3.2 *Let $\omega \in \{0, 1\}^{\mathbb{N}}$ be a sequence satisfying property (l, L) , then ω is not a central sequence.*

Proof. First of all we prove that if ω' is proximal to ω and ω' is uniformly recurrent then ω' is the zeros sequence. For, let us assume that ω' has aforementioned properties and take a block $(\omega'_0, \omega'_1, \dots, \omega'_{m-1})$. Then there exists l such that this block is contained in any continuous subblock of length l of ω' . But ω is proximal to ω' , thus for any $L > 0$ there exists $n \geq 0$ such that $\{\omega\}_n^{n+L-1} = \{\omega'\}_n^{n+L-1}$. Let us choose $L = L(l)$ such that any subblock of ω of length L contains l successive zeros.

As a result of our choices, the block $(\omega'_0, \omega'_1, \dots, \omega'_{m-1})$ is a subblock of any subblock of $\{\omega\}_n^{n+L-1}$ of length l , in particular the zeros block of length l .

Thus we have proved that ω is proximal to only one uniform recurrent sequence, namely the zeros sequence.

Suppose, contrary to the assertion of the lemma that ω is a central sequence. Then there exists V , a neighborhood of the zeros sequence, such that $\omega_n = 1$ iff $T^n \omega \in V$. V contains an open set which contains the zeros sequence, thus there exists $l \geq 1$, such that any word x that begins with l zeros is inside V (l can not be zero, because then ω is ones's sequence which does not satisfy (l, L) property); ω is proximal to the zeros sequence, thus there exists $k \geq 0$ such that $\{\omega\}_k^{k+l-1}$ is the zeros block. But in this case $T^k \omega \in V$ and therefore $\omega_k = 1$.

Thus we have got a contradiction to our assumption that ω is central.

□

Theorem 2.3.1 *Let (X, \mathbb{B}, μ, T) be a weak mixing invertible m.p.s.. Then there exist a symbolic weak mixing system (Y, \mathbb{B}, ν, T) and $y_0 \in \{0, 1\}^{\mathbb{N}}$ a generic point of Y with $0 < d(y_0) < 1$ (where d is the density of ones) such that every sequence $y \leq y_0$ (for every n we have $y(n) \leq y_0(n)$) is not a central sequence and Y is a factor of X .*

Proof. By remark 2.3.2 the system (X, \mathbb{B}, μ, T) satisfies all the requirements of lemma 2.3.1 and therefore for any $n > 1$ there exists $C_n \in \mathbb{B}$ such that $X = \bigcup_{i=0}^n T^i C_n$ with $\frac{1}{n} \leq \mu(C_n) \leq \frac{1}{n-1}$, and every point inside C_n returns back to C_n by at most $n+1$ iterations of T .

Now we construct $A \in \mathbb{B}$ of a positive measure which is bounded by a predefined number α by the following procedure.

Let us choose $\{L_l\}$ a sequence of positive natural numbers (for every l we assume that $L_l \geq 2$) such that $\sum_{l=1}^{\infty} \frac{l}{L_l-1} = \alpha$. Then for any L_l let us take $C_{L_l} \in \mathbb{B}$ as above with $X = \bigcup_{i=0}^{L_l} T^i C_{L_l}$ and, finally, take

$$A = \bigcup_{l=1}^{\infty} \left(\bigcup_{i=0}^{L_l-1} T^i C_{L_l} \right)$$

The following estimations on the measure of A are obvious:

$$\frac{1}{L_1} \leq \mu(A) \leq \sum_{l=1}^{\infty} l \mu(C_{L_l}) \leq \sum_{l=1}^{\infty} \frac{l}{L_l-1} = \alpha$$

Let $\phi_A : X \rightarrow \{0,1\}^{\mathbb{N}}$ to be defined by $\omega = \phi_A(x)$ iff $\omega(n) = 1 - 1_A(T^n x)$. Then obviously ϕ_A is measurable and $\phi_A \circ T = T \circ \phi_A$ (where T on the right hand is the usual shift transformation). Let us define $Y \doteq \phi_A(X)$, then $(Y, \mathbb{B}_Y, (\phi_A)_* \mu, T)$ is a m.p.s. and a factor of X , thus is a weak mixing system (for any $B \in \mathbb{B}_Y$ we define $(\phi_A)_* \mu \doteq \mu(\phi_A^{-1}(B))$).

Let us denote by X' the following subset of X

$$X' \doteq \bigcap_{l=1}^{\infty} \left(\bigcup_{i=0}^{L_l+1} T^i C_{L_l} \right)$$

It is obvious that $\mu(X') = \mu(X)$. Let $G \subset X$ be the set of generic points in X (X is a compact metric space and therefore the notion of a generic point is well defined). By the ergodic theorem $\mu(G) = \mu(X)$, and therefore $\mu(G \cap X') = \mu(X)$ and thus the measure of $\phi_A(G \cap X')$ in Y is equal to the measure of whole Y . But (Y, T) is ergodic (even weak mixing) therefore almost every point of $\phi_A(G \cap X')$ is generic. Choose $y_0 \in \phi_A(G \cap X')$ to be generic in Y . Then there exists $x_0 \in G \cap X'$ such that $y_0 = \phi_A(x_0)$ (by using the ergodic theorem once again, we can add one more condition on x_0 , namely $\frac{1}{N} \sum_{n=1}^N 1_A(T^n x_0) \rightarrow \mu(A)$ [it should be done because 1_A might be a non continuous function]). It is obvious that $d(y_0) = \mu(A^c)$ and thus $d(y_0) \geq 1 - \alpha$. By the choice of A and x_0 it follows that y_0 satisfies (l, L) property. Every sequence $y \leq y_0$ is again satisfies (l, L) property and, thus, by lemma 2.3.2, it follows that y is not a central sequence.

□

Remark 2.3.3 *For any $0 < \alpha < 1$ we can construct y_0 in the formulation of the theorem with $d(y_0) \geq \alpha$.*

By combining theorem 2.3.1 with remark 2.3.3 we obtain the following statement.

Theorem 2.3.2 *For any $0 < \alpha < 1$ there exists $A \in \mathbb{N}$ a WM set with $d(y_0) \geq \alpha$ which is not central and such that no subset B of A is a central sequence.*

3 Solvability of linear diophantine equations within WM sets

3.1 Proof of Sufficiency

We restate the main result of this section which was formulated in section 1.4.1.

Theorem 1.4.1 *Let $B \in \mathbb{Q}^{r \times k}$ and $\vec{d} \in \mathbb{Q}^r$. The system of linear equations*

$$B\vec{x} = \vec{d} \quad (3.1)$$

is solvable within every WM set \Leftrightarrow there exist two vectors $\vec{x}_1 = (a_1, a_2, \dots, a_k)^t$, $\vec{x}_2 = (b_1, b_2, \dots, b_k)^t \in \mathbb{N}^k$, disjoint sets $E, F_1, \dots, F_l \subset \{1, 2, \dots, k\}$, $E \cup F_1 \cup \dots \cup F_l = \{1, 2, \dots, k\}$, such that:

a) for every $i, j \in E$, $i \neq j$

$$\det \begin{pmatrix} a_i & b_i \\ a_j & b_j \end{pmatrix} \neq 0.$$

b) for every $p \in \{1, \dots, l\}$ there exist $c_1^p, c_2^p \in \mathbb{N}$, such that for every $i \in F_p$ we have $a_i = c_1^p$, $b_i = c_2^p$ and for every $j \in \{1, \dots, k\} \setminus F_p$ we require

$$\det \begin{pmatrix} a_j & b_j \\ c_1^p & c_2^p \end{pmatrix} \neq 0.$$

c) There exist $f^1, \dots, f^l \in \mathbb{Z}$ such that setting $f_i = f^p$ for $p \in \{1, \dots, l\}$ and $i \in F_p$, then the affine space of solutions of a system $B\vec{x} = \vec{d}$ contains

$$\{(a_1n + b_1m + f_1, \dots, a_kn + b_km + f_k)^t \mid n, m \in \mathbb{N}\}.$$

Notation: *We introduce the scalar product of two vectors v, w of the length N as follows:*

$$\langle v, w \rangle_N \doteq \frac{1}{N} \sum_{n=1}^N v(n)w(n).$$

We denote by $L^2(N)$ the Hilbert space of all real vectors of the length N with the aforementioned scalar product.

We define: $\|w\|_N^2 \doteq \langle w, w \rangle_N$.

First we state the following proposition which is a very useful tool in the proof of the sufficiency of the conditions of theorem 1.4.1.

Proposition 3.1.1 *Let $A_i \subset \mathbb{N}$ ($1 \leq i \leq k$) be WM sets. Let*

$\xi_i(n) \doteq 1_{A_i}(n) - d(A_i)$, where $d(A_i)$ denotes density of A_i . Suppose there are $(a_1, b_1), (a_2, b_2), \dots, (a_k, b_k) \in (\mathbb{Z} \setminus \{0\})^2$, such that $a_i > 0$, $1 \leq i \leq k$, and for every $i \neq j$

$$\det \begin{pmatrix} a_i & b_i \\ a_j & b_j \end{pmatrix} \neq 0.$$

Then for every $\varepsilon > 0$ there exists $M(\varepsilon) \in \mathbb{N}$, such that for every $M \geq M(\varepsilon)$ there exists $N(M, \varepsilon) \in \mathbb{N}$, such that for every $N \geq N(M, \varepsilon)$

$$\|w\|_N < \varepsilon,$$

where $w(n) \doteq \frac{1}{M} \sum_{m=1}^M \xi_1(a_1n + b_1m) \xi_2(a_2n + b_2m) \dots \xi_k(a_kn + b_km)$ for every $n = 1, 2, \dots, N$.

Since the proof of proposition 3.1.1 involves many technical details, first we show how our main result follows from it. Afterwards we state and prove all the lemmas necessary for a proof of proposition 3.1.1 and define all the required concepts.

We will need an easy consequence of proposition 3.1.1.

Corollary 3.1.1 *Let A be a WM set. Let $k \in \mathbb{N}$, suppose $(a_1, b_1), (a_2, b_2), \dots, (a_k, b_k) \in (\mathbb{Z} \setminus \{0\})^2$ satisfy all requirements of proposition 3.1.1 and suppose $f_1, \dots, f_k \in \mathbb{Z}$. Then for every $\delta > 0$ there exists $M(\delta)$ such that $\forall M \geq M(\delta)$ there exists $N(M, \delta)$ such that $\forall N \geq N(M, \delta)$ we have*

$$|\|v\|_N - d^k(A)| < \delta,$$

where $v(n) \doteq \frac{1}{M} \sum_{m=1}^M 1_A(a_1n + b_1m + f_1) 1_A(a_2n + b_2m + f_2) \dots 1_A(a_kn + b_km + f_k)$ for every $n = 1, 2, \dots, N$.

Proof. We can write $v(n)$ in the following form:

$$v(n) = \frac{1}{M} \sum_{m=1}^M (\xi(a_1n + b_1m) + d(A)) (\xi(a_2n + b_2m) + d(A)) \dots (\xi(a_kn + b_km) + d(A)),$$

for every $n = 1, 2, \dots, N$. We again introduce normalized WM sequences $\xi_i(n) = \xi(n + f_i)$. Then by use of triangular inequality and proposition 3.1.1 it follows that for big enough M and N (which depends on M) $\|v\|_N$ is as close as we wish to $d^k(A)$. The latter finishes the proof.

□

Proof. (of the theorem 1.4.1, \Leftarrow)

By corollary 3.1.1 it follows that the vector v defined by

$$v(n) \doteq \frac{1}{M} \sum_{m=1}^M 1_A(a_1n + b_1m + f_1) 1_A(a_2n + b_2m + f_2) \dots 1_A(a_kn + b_km + f_k),$$

for every $n = 1, 2, \dots, N$ is not identically zero for big enough M and N . The latter is possible only if for some $n, m \in \mathbb{N}$ we have

$$(a_1n + b_1m + f_1, a_2n + b_2m + f_2, \dots, a_kn + b_km + f_k) \in A^k.$$

□

Now we state and prove all the claims that are required in order to prove proposition 3.1.1.

Definition 3.1.1 Let ξ be a WM-sequence of zero average. The autocorrelation function of ξ of the length $j \in \mathbb{N}$ with the shifts $\{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}\}$ (all shifts are integers) is the sequence $\psi_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^j$ which is defined as follows: for $j > 1$

$$\psi_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^j(n) = \psi_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_{j-1}, i_{j-1}\}}^{j-1}(n + r_j) \psi_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_{j-1}, i_{j-1}\}}^{j-1}(n + r_j + i_j),$$

for $j = 1$ the autocorrelation function is defined as

$$\psi_{\{r_1, i_1\}}^1(n) = \xi(n + r_1) \xi(n + r_1 + i_1).$$

Remark 3.1.1 For any sequence ψ we define $\psi(-n) = 0$ for every $n \in \mathbb{N}$.

Lemma 3.1.1 Let ξ be a WM-sequence of zero average and suppose $\varepsilon, \delta > 0$. Then for every $j \geq 1$, $\{c_1, c_2, \dots, c_j\} \in (\mathbb{Z} \setminus \{0\})^j$ and $\{r_1, r_2, \dots, r_j\} \in (\mathbb{Z})^j$ there exists $I = I(\varepsilon, \delta, c_1, \dots, c_n)$, such that there exists a set $S \subset [-I, I]^j$ of density at least $1 - \delta$ and there exists $N(I, \varepsilon) \in \mathbb{N}$, such that for every $N \geq N(I, \varepsilon)$ there exists $L(N, I, \varepsilon)$ such that for every $L \geq L(N, I, \varepsilon)$

$$\frac{1}{L} \sum_{l=1}^L \left(\frac{1}{N} \sum_{n=1}^N \psi_{\{r_1, c_1 i_1\}, \{r_2, c_2 i_2\}, \dots, \{r_j, c_j i_j\}}^j(l + bn) \right)^2 < \varepsilon,$$

for every $\{i_1, i_2, \dots, i_j\} \in S$.

Proof. We note that it is sufficient to prove the lemma in the case $c_1 = c_2 = \dots = c_j = 1$, since if the average of nonnegative numbers over a complete lattice is small, then the average over a sublattice of a fixed positive density is also small. Recall that $\xi \in X_\xi \doteq \{T^n \xi\}_{n=0}^\infty \subset \text{supp}(\xi)^\mathbb{N}$, where T is a usual shift to the left on the dynamical system $\text{supp}(\xi)^\mathbb{N}$, and by the assumption that ξ is a WM-sequence of zero average it follows that ξ is a generic point of the weak-mixing system $(X_\xi, \mathbb{B}_{X_\xi}, \mu, T)$ and the function $f : f(\omega) \doteq \omega_0$ has zero integral.

We define functions $g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}$ on X_ξ inductively. Let $g_\emptyset \doteq f$. Define $g_{\{r_1, i_1\}, \dots, \{r_{j-1}, i_{j-1}\}, \{r_j, i_j\}} \doteq T^{r_j} (g_{\{r_1, i_1\}, \dots, \{r_{j-1}, i_{j-1}\}} T^{i_j} g_{\{r_1, i_1\}, \dots, \{r_{j-1}, i_{j-1}\}})$. Define the functions $g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^* = \prod_{\epsilon \in V_j^*} f \circ T^{r_1 + \dots + r_j + \epsilon_1 i_1 + \dots + \epsilon_j i_j}$, where V_j is a j -dimensional discrete cube $\{0, 1\}^j$ and V_j^* is the whole j -dimensional discrete cube except the zero point. (Note that $g = (T^{r_1 + \dots + r_j} \circ f) g^*$, where we have omitted subscripts.)

The following has been proven by Host and Kra in [12] (theorem 13.1):

Let (X, μ, T) be an ergodic system. Given an integer k and 2^k bounded functions f_ϵ on X , $\epsilon \in V_k$, the functions

$$\prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1) \times \dots \times [M_k, N_k)} \prod_{\epsilon \in V_k^*} f_\epsilon \circ T^{\epsilon_1 n_1 + \dots + \epsilon_k n_k}$$

converge in $L^2(\mu)$ to the limit function

$$\mathbb{E} \left(\bigotimes_{\epsilon \in V_k^*} f_\epsilon | \tau^{[k]*} \right) (x),$$

when $N_1 - M_1, \dots, N_k - M_k$ tend to $+\infty$. The σ -algebra $\tau^{[k]*}$ is identified with the so-called characteristic factor $Z_{k-1}(X)$.

The characteristic factors $Z_k(X)$ are defined for arbitrary ergodic systems, and what is important for our purposes is that in our case of the weak-mixing system X_ξ , all the factors $Z_{k-1}(X_\xi)$ are trivial.

Therefore, the limit function in our case will be a constant. By integrating the limit function we obtain that this constant is equal to $\prod_{\epsilon \in V_k^*} \int_{X_\xi} f_\epsilon d\mu$.

From the theorem of Host and Kra, applied to the weak-mixing system $X_\xi \times X_\xi$ and the functions $f_\epsilon(x) = f \otimes f$ for every $\epsilon \in V_k$, we obtain for every Folner sequence $\{F_n\}$ in \mathbb{N}^j that an average over the multi-index $\{i_1, \dots, i_j\}$ of $g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^* \otimes g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^*$ on F_n 's converges to zero (the integral of $f \otimes f$ is zero). If we would take another Folner sequence $\{G_n\}$ in \mathbb{N}^j then for the same $\{r_1, \dots, r_j\}$ the closeness of an average of $g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^* \otimes g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^*$ on G_n to zero depends only the size of the box G_n . Namely, if all edges of a box are big enough then the aforementioned average is small.

As a result we have

For every $\varepsilon > 0$, $j \in \mathbb{N}$ and every fixed $\{r_1, r_2, \dots, r_j\} \in \mathbb{N}^j$, there exists a subset $R \subset \mathbb{N}^j$ with lower density equal to one, such that

$$\left(\int_{X_\xi} g_{\{r_1, i_1\}, \dots, \{r_{j-1}, i_{j-1}\}, \{r_j, i_j\}} d\mu \right)^2 < \varepsilon, \quad (3.2)$$

for every $\{i_1, i_2, \dots, i_j\} \in R$.

We note that $\psi_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^j(l + bn) = g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}(T^{l+bn}\xi)$.

The definition of the sequences ψ^j implies

$$\begin{aligned} & \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L \left(\frac{1}{N} \sum_{n=1}^N \psi_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^j(l + bn) \right)^2 \\ &= \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L \left(\frac{1}{N} \sum_{n=1}^N \psi_{\{\pm r_1, \pm i_1\}, \{\pm r_2, \pm i_2\}, \dots, \{\pm r_j, \pm i_j\}}^j(l \pm bn) \right)^2. \end{aligned}$$

Therefore, in order to prove the Lemma 3.1.1 it is sufficient to show the following:

For every $\varepsilon, \delta > 0$ and for a priori chosen $r_1, r_2, \dots, r_j, b \in \mathbb{N}$ there exists $I(\varepsilon, \delta) \in \mathbb{N}$, such that for every $I \geq I(\varepsilon, \delta)$ there exists a subset $S \subset [1, I]^j$ of density at least $1 - \delta$ (namely, we have $\frac{|S \cap [1, I]^j|}{I^j} \geq 1 - \delta$) and there exists $N(I, \varepsilon) \in \mathbb{N}$, such that for every $N \geq N(I, \varepsilon)$ there exists $L(N, I, \varepsilon) \in \mathbb{N}$ such that for every $L \geq L(N, I, \varepsilon)$ the following holds for every $\{i_1, i_2, \dots, i_j\} \in S$:

$$\frac{1}{L} \sum_{l=1}^L \left(\frac{1}{N} \sum_{n=1}^N \psi_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^j(l + bn) \right)^2 < \varepsilon.$$

Assume that $r_1, r_2, \dots, r_j, b \in \mathbb{N}$. Continuity of the function $g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}$ and the genericity of the point $\xi \in X_\xi$ yields

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L \left(\frac{1}{N} \sum_{n=1}^N \psi_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^j(l + bn) \right)^2$$

$$\begin{aligned}
&= \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L \left(\frac{1}{N} \sum_{n=1}^N T^{bn} g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}} (T^l \xi) \right)^2 \\
&= \int_{X_\xi} \left(\frac{1}{N} \sum_{n=1}^N T^{bn} g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}} \right)^2 d\mu.
\end{aligned} \tag{3.3}$$

By combining the ergodic theorem, applied to the weak-mixing system $(X_\xi, \mathbb{B}, \mu, T^b)$, with disjointness of any weak-mixing system from the cyclic system on b elements we note that

$$\frac{1}{N} \sum_{n=1}^N T^{bn} g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}} \xrightarrow{N \rightarrow \infty} \int_{X_\xi} g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}} d\mu. \tag{3.4}$$

From (3.2) there exists $I(\varepsilon, \delta) \in \mathbb{N}$ big enough, such that for every $I \geq I(\varepsilon, \delta)$ there exists a set $S \subset [1, I]^j$ of density at least $1 - \delta$ such that

$$\left(\int_{X_\xi} g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}} d\mu \right)^2 < \frac{\varepsilon}{4},$$

for all $\{i_1, i_2, \dots, i_j\} \in S$.

From equation (3.4) follows that there exists $N(I, \varepsilon) \in \mathbb{N}$, such that for every $N \geq N(I, \varepsilon)$ we have

$$\int_{X_\xi} \left(\frac{1}{N} \sum_{n=1}^N T^{bn} g_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}} \right)^2 < \frac{\varepsilon}{2},$$

for all $\{i_1, i_2, \dots, i_j\} \in S$.

Finally, equation (3.3) implies that there exists $L(N, I, \varepsilon) \in \mathbb{N}$, such that for every $L \geq L(N, I, \varepsilon)$ we obtain

$$\frac{1}{L} \sum_{l=1}^L \left(\frac{1}{N} \sum_{n=1}^N \psi_{\{r_1, i_1\}, \{r_2, i_2\}, \dots, \{r_j, i_j\}}^j (l + bn) \right)^2 < \varepsilon,$$

for all $\{i_1, i_2, \dots, i_j\} \in S$.

□

The following lemma is a generalization of the previous lemma for a product of several autocorrelation functions.

Lemma 3.1.2 *Let $\psi_{\{r_1^1, i_1\}, \{r_2^1, i_2\}, \dots, \{r_j^1, i_j\}}^{1,j}, \dots, \psi_{\{r_1^k, i_1\}, \{r_2^k, i_2\}, \dots, \{r_j^k, i_j\}}^{k,j}$ be autocorrelation functions of length j of WM-sequences ξ_1, \dots, ξ_k of zero average, $\{c_1^1, \dots, c_j^1, \dots, c_1^k, \dots, c_j^k\} \in (\mathbb{Z} \setminus \{0\})^{jk}$ and $\varepsilon, \delta > 0$. Suppose $(a_1, b_1), (a_2, b_2), \dots, (a_k, b_k) \in \mathbb{Z}^2$, such that $a_i > 0, b_i \neq 0, 1 \leq i \leq k$ and for every $i \neq j$*

$$\det \begin{pmatrix} a_i & b_i \\ a_j & b_j \end{pmatrix} \neq 0.$$

Then there exists $I(\varepsilon, \delta) \in \mathbb{N}$, such that for every $I \geq I(\varepsilon, \delta)$ there exist $S \subset [-I, I]^j$ of density at least $1 - \delta$, $M(I, \varepsilon) \in \mathbb{N}$, such that for every $M \geq M(I, \varepsilon)$ there exists $X(M, I, \varepsilon) \in \mathbb{N}$, such that for every $X \geq X(M, I, \varepsilon)$

$$\frac{1}{X} \sum_{x=1}^X \left(\frac{1}{M} \sum_{m=1}^M \psi_{\{r_1^1, c_1^1 i_1\}, \{r_2^1, c_2^1 i_2\}, \dots, \{r_j^1, c_j^1 i_j\}}^{1,j} (a_1 x + b_1 m) \dots \right.$$

$$\left. \psi_{\{r_1^k, c_1^k i_1\}, \{r_2^k, c_2^k i_2\}, \dots, \{r_j^k, c_j^k i_j\}}^{k,j} (a_k x + b_k m) \right)^2 < \varepsilon,$$

for every $\{i_1, i_2, \dots, i_j\} \in S$.

Proof. The proof is by induction on k . The case $k = 1$ (and arbitrary j) follows from the Lemma 3.1.1 and the Proposition 6.1.

Suppose that the statement holds for $k - 1$.

Denote by

$$v_m(x) \doteq \psi_{\{r_1^1, c_1^1 i_1\}, \dots, \{r_j^1, c_j^1 i_j\}}^{1,j} (a_1 x + b_1 m) \dots \psi_{\{r_1^k, c_1^k i_1\}, \dots, \{r_j^k, c_j^k i_j\}}^{k,j} (a_k x + b_k m).$$

The van der Corput lemma (lemma 6.1 of the appendix) implies that it is sufficient to show the existence of $\mathbb{I}(\varepsilon, \delta) \in \mathbb{N}$, such that for every $\mathbb{I} \geq \mathbb{I}(\varepsilon, \delta)$ there exists a set $S \subset [-\mathbb{I}, \mathbb{I}]^j$ of density at least $1 - \delta$ and there exists $I(\varepsilon, \mathbb{I})$ big enough ($I(\varepsilon, \mathbb{I}) \geq I'(\varepsilon)$ from van der Corput Lemma), such that for most of the i 's in the interval $\{1, 2, \dots, I(\varepsilon, \mathbb{I})\}$ (density of such i 's should be at least $1 - \frac{\varepsilon}{3}$) there exists $M(I(\varepsilon, \mathbb{I}), \mathbb{I}, \varepsilon) \in \mathbb{N}$, such that for every $M \geq M(I(\varepsilon, \mathbb{I}), \mathbb{I}, \varepsilon)$

$$\left| \frac{1}{M} \sum_{m=1}^M < v_m, v_{m+i} >_X \right| < \frac{\varepsilon}{2}, \quad (3.5)$$

for all $\{i_1, \dots, i_j\} \in S$.

In our case we obtain

$$\begin{aligned} & \left| \frac{1}{M} \sum_{m=1}^M < v_m, v_{m+i} >_X \right| = \\ & \left| \frac{1}{X} \sum_{x=1}^X \frac{1}{M} \sum_{m=1}^M \psi_{\{r_1^1, c_1^1 i_1\}, \dots, \{r_j^1, c_j^1 i_j\}, \{0, b_1 i\}}^{1,j+1} (a_1 x + b_1 m) \dots \right. \\ & \left. \psi_{\{r_1^k, c_1^k i_1\}, \dots, \{r_j^k, c_j^k i_j\}, \{0, b_k i\}}^{k,j+1} (a_k x + b_k m) \right| = \tilde{A}. \end{aligned}$$

Denote $y = a_1 x + b_1 m$. Assume that $(a_1, b_1) = d$. Denote

$$\tilde{B}_{y,m} = \psi_{\{r_1^1, c_1^1 i_1\}, \dots, \{r_j^1, c_j^1 i_j\}, \{0, b_1 i\}}^{1,j+1} (y) \dots \psi_{\{r_1^k, c_1^k i_1\}, \dots, \{r_j^k, c_j^k i_j\}, \{0, b_k i\}}^{k,j+1} (a'_k y + b'_k m),$$

where $a'_p = \frac{a_p}{a_1}$, $b'_p = b_p - a'_p b_1$. Now we rewrite \tilde{A} in the following way

$$\tilde{A} = \left| a_1 \frac{1}{Y} \left(\sum_{l=0}^{\frac{a_1}{d}-1} \sum_{y \equiv dl \pmod{a_1}}^Y \frac{1}{M} \sum_{m \equiv \phi(l) \pmod{\frac{a_1}{d}}}^M \tilde{B}_{y,m} \right) \right| + \delta_{X,M}.$$

Here ϕ is the one to one function from $\mathbb{Z}_{\frac{a_1}{d}}$ onto itself, such that $\phi(l)\frac{b_1}{d} \equiv l \pmod{\frac{a_1}{d}}$ for every $0 \leq l \leq \frac{a_1}{d} - 1$, $Y = a_1 X$, a'_p, b'_p as above and $\delta_{X,M}$ accounts for the fact that for small y 's and y 's close to Y there is a difference between elements that are taken in the expression for \tilde{A} and in the expression on the right hand side of the last equation. Nevertheless, we have $\delta_{X,M} \rightarrow 0$ if $\frac{M}{X} \rightarrow 0$.

It will suffice to prove (Cauchy-Schwartz inequality) that there exists $\mathbb{I}(\varepsilon, \delta) \in \mathbb{N}$, such that for every $\mathbb{I} \geq \mathbb{I}(\varepsilon, \delta)$ there exists a set $S \subset [-\mathbb{I}, \mathbb{I}]^j$ of density at least $1 - \delta$ and there exist $I(\varepsilon, \mathbb{I}) \in \mathbb{N}$, $M(I(\varepsilon, \mathbb{I})) \in \mathbb{N}$, such that for every $M \geq M(I(\varepsilon, \mathbb{I}))$ there exists $X(M, \varepsilon) \in \mathbb{N}$ such that for every $X \geq X(M, \varepsilon)$, and for a set of i 's in the interval $\{1, 2, \dots, I(\varepsilon, \mathbb{I})\}$ of density $1 - \frac{\varepsilon}{3}$ we have

$$a_1 \frac{1}{Y} \sum_{y \equiv dl \pmod{a_1}}^Y \left(\frac{1}{M} \sum_{m \equiv \phi(l) \pmod{\frac{a_1}{d}}}^M \tilde{C}_{y,m} \right)^2 < \left(\frac{\varepsilon d}{3a_1} \right)^2, \quad (3.6)$$

for all $0 \leq l \leq \frac{a_1}{d} - 1$, where

$$\begin{aligned} \tilde{C}_{y,m} &= \psi_{\{r_1^2, c_1^2 i_1\}, \dots, \{r_j^2, c_j^2 i_j\}, \{0, b_2 i\}}^{2,j+1} (a'_2 y + b'_2 m) \dots \\ &\quad \psi_{\{r_1^k, c_1^k i_1\}, \dots, \{r_j^k, c_j^k i_j\}, \{0, b_k i\}}^{k,j+1} (a'_k y + b'_k m). \end{aligned}$$

We rewrite the inequality (3.6) for a fixed l as follows:

Denote z and n , such that $y = za_1 + dl$ and $m = n\frac{a_1}{d} + \phi(l)$. As a result we obtain

$$\begin{aligned} \frac{1}{Z} \sum_{z=1}^Z \left(\frac{d}{Na_1} \sum_{n=1}^N \psi_{sh_2}^{2,j+1} (t_{n,z,l}^2) \dots \psi_{sh_k}^{k,j+1} (t_{n,z,l}^k) \right)^2 = \\ \frac{1}{Z} \sum_{z=1}^Z \left(\frac{d}{Na_1} \sum_{n=1}^N \psi_{sh_2}^{2,j+1} (a_2 z + c_2 n + r_2) \dots \psi_{sh_k}^{k,j+1} (a_k z + c_k n + r_k) \right)^2 \doteq \tilde{D}, \end{aligned}$$

where $sh_p = \{\{r_1^p, c_1^p i_1\}, \dots, \{r_j^p, c_j^p i_j\}, \{0, b_p i\}\}$,

$$t_{n,z,l}^p = \frac{a_p(a_1 z + dl) + (a_1 b_p - a_p b_1)(\frac{a_1}{d} n + \phi(l))}{a_1}, \quad r_p = \frac{a_p l + (a_1 b_p - a_p b_1)\phi(l)}{a_1},$$

$c_p = \frac{a_1 b_p - a_p b_1}{d} \neq 0$, $Z = \frac{Y}{a_1}$ and $N = \frac{Md}{a_1}$. The expression $\frac{a_p l + (a_1 b_p - a_p b_1)\phi(l)}{a_1} \in \mathbb{Z}$ (from the condition on the function ϕ).

From the conditions of the lemma we obtain for every $p \neq q$, $p, q > 1$

$$\det \begin{pmatrix} a_p & c_p \\ a_q & c_q \end{pmatrix} = \frac{a_1 \det \begin{pmatrix} a_p & b_p \\ a_q & b_q \end{pmatrix}}{d} \neq 0.$$

Therefore, we have \tilde{D} can be rewritten

$$\begin{aligned} \tilde{D} &= \frac{1}{Z} \sum_{z=1}^Z \left(\frac{1}{Na_1} \sum_{n=1}^N \psi_{\{r_1^2, c_1^2 i_1\}, \dots, \{r_j^2, c_j^2 i_j\}, \{r_2, b_2 i\}}^{2,j+1} (a_2 z + c_2 n) \dots \right. \\ &\quad \left. \psi_{\{r_1^k, c_1^k i_1\}, \dots, \{r_j^k, c_j^k i_j\}, \{r_k, b_k i\}}^{k,j+1} (a_k z + c_k n) \right)^2. \end{aligned}$$

By the induction hypothesis there exists $\mathbb{I}(\varepsilon, \delta) \in \mathbb{N}$ big enough, such that for every $\mathbb{I} \geq \mathbb{I}(\varepsilon, \delta)$ there exist a subset $S \subset [-\mathbb{I}, \mathbb{I}]^{j+1}$ of density at least $1 - \delta^2$ and $N(\mathbb{I}, \varepsilon) \in \mathbb{N}$, such that for every $N \geq N(\mathbb{I}, \varepsilon)$ there exists $Z(N, \mathbb{I}, \varepsilon) \in \mathbb{N}$, such that for every $Z \geq Z(N, \mathbb{I}, \varepsilon)$

$$\tilde{D} < \left(\frac{\varepsilon}{3a_1} \right)^2, \quad (3.7)$$

for all $\{i_1, \dots, i_j, i\} \in S$.

For every $(i_1, \dots, i_j) \in [-\mathbb{I}, \mathbb{I}]^j$ we denote by S_{i_1, \dots, i_j} the following subset of $[-\mathbb{I}, \mathbb{I}]$

$$S_{i_1, \dots, i_j} = \{i \in [-\mathbb{I}, \mathbb{I}] \mid (i_1, \dots, i_j, i) \in S\}.$$

Then there exists a set $T \subset [-\mathbb{I}, \mathbb{I}]^j$ of density at least $1 - \delta$, such that for every $(i_1, \dots, i_j) \in T$ the density of S_{i_1, \dots, i_j} is at least $1 - \delta$. Let $\delta < \frac{\varepsilon}{7}$ and $\mathbb{I} > \max_l (\max(I'(\varepsilon), \mathbb{I}(\varepsilon, \delta)))$ ($I'(\varepsilon)$ is taken from van der Corput lemma). By taking $N(\mathbb{I}, \varepsilon, \delta)$, follows from the inequality (3.7) that there exists $M(\mathbb{I}, \varepsilon, \delta) \in \mathbb{N}$, such that for every $M \geq M(\mathbb{I}, \varepsilon, \delta)$ there exists $X(M, \mathbb{I}, \varepsilon, \delta) \in \mathbb{N}$, such that for every $X \geq X(M, \mathbb{I}, \varepsilon, \delta)$ the inequality (3.5) holds for every fixed $(i_1, \dots, i_j) \in T$ for a set of i 's within the interval $\{1, \dots, \mathbb{I}\}$ of density at least $1 - \frac{\varepsilon}{3}$. The lemma follows from the van der Corput lemma. \square

Proof. (of Proposition 3.1.1)

Denote by $v_m(x) \doteq \xi_1(a_1x + b_1m) \dots \xi_k(a_kx + b_km)$. Then for every $i \in \mathbb{N}$

$$\left| \frac{1}{M} \sum_{m=1}^M \langle v_m, v_{m+i} \rangle_X \right| =$$

$$\left| \frac{1}{X} \sum_{x=1}^X \frac{1}{M} \sum_{m=1}^M \psi_{\{0, b_1 i\}}^{1,1}(a_1x + b_1m) \dots \psi_{\{0, b_k i\}}^{k,1}(a_kx + b_km) \right| \doteq \tilde{A},$$

where functions $\psi^{p,j}$'s are autocorrelation functions of ξ_p 's of the length j .

Again, as in the proof of the lemma 3.1.2 we denote $y = a_1x + b_1m$. We proceed with the analysis of the expression (\tilde{A}) and by the same technique which was used in the proof of the lemma 3.1.2 we conclude the following:

In order to prove that $\tilde{A} < \frac{\varepsilon}{2}$ for a set of i 's within the appropriate interval $\{1, 2, \dots, I(\varepsilon)\}$ it is sufficient to prove that there exists $I(\varepsilon) \in \mathbb{N}$ big enough and $N(I(\varepsilon), \varepsilon) \in \mathbb{N}$, such that for every $N \geq N(I(\varepsilon), \varepsilon)$ there exists $Z(N, \varepsilon) \in \mathbb{N}$, such that for every $Z \geq Z(N, \varepsilon)$

$$\frac{1}{Z} \sum_{z=1}^Z \left(\frac{1}{Na_1} \sum_{n=1}^N \psi_{\{r_2, b_2 i\}}^{2,1}(a_2z + c_2n) \dots \psi_{\{r_k, b_k i\}}^{k,1}(a_kz + c_kn) \right)^2 < \left(\frac{\varepsilon}{3a_1} \right)^2,$$

for a set i 's within the interval $\{1, \dots, I(\varepsilon)\}$ of density $1 - \frac{\varepsilon}{3}$.

The last statement follows from lemma 3.1.2. The proposition follows from the van der Corput lemma. \square

3.2 Probabilistic constructions of WM sets

The goal of this section is to prove the necessity of the conditions of theorem 1.4.1 and the following proposition is the main tool for this task.

Proposition 3.2.1 *Let $a, b \in \mathbb{N}$, $c \in \mathbb{Z}$ such that $a \neq b$. Then there exists a WM set A such that within it the equation*

$$ax = by + c \quad (3.8)$$

is unsolvable, i.e., for every $(x, y) \in A^2$ we have $ax \neq by + c$.

Remark 3.2.1 The proposition is a particular case of theorem 1.4.1. It is a crucial ingredient in proving the necessity direction of the theorem in general.

Proof. Let $S \subset \mathbb{N}$. We construct from S a new set A_S , such that within it the equation $ax = by + c$ is unsolvable. Without loss of generality, suppose that $a < b$.

Assume $(a, b) = 1$ (the general case follows easily). The equation $ax = by + c$ is solvable only if $x \equiv \phi(a, b, c) \pmod{b}$, where $\phi(a, b, c) : 0 \leq \phi(a, b, c) < b$ is determined uniquely (if the equation has a solution at all, otherwise any WM set will provide an example). Let us denote $l_0 \doteq \phi(a, b, c)$. We define inductively a sequence $\{l_i\} \subset \mathbb{N} \cup \{0\}$. If a pair (x, y) is a solution of the equation and $y \in b^i \mathbb{N} + l_{i-1}$ then there exists $l_i \in \{0, 1, \dots, b^{i+1} - 1\}$ such that $x \in b^{i+1} \mathbb{N} + l_i$.

We define the sets $H_i \doteq b^i \mathbb{N} + l_{i-1}; i \in \mathbb{N}$. We prove that for every $i \in \mathbb{N} : H_{i+1} \subset H_i$. All elements of H_{i+1} are in the same class modulo b^i . So, if we show for some $x \in H_{i+1}$ that $x \equiv l_{i-1} \pmod{b^{i-1}}$ then we are done. For $i = 1$ we know that if $y \in \mathbb{N}$ then $x : ax = by + c$ has to be in H_1 . Therefore for $x \in H_2$ such that there exists $y \in H_1$ such that $ax = by + c$ we have that $x \in H_1$. Therefore, we have shown that $H_2 \subset H_1$. For $i > 1$ there exists $x \in H_{i+1}$ such that there exists $y \in H_i$ with $ax = by + c$. By induction $H_i \subset H_{i-1}$. Therefore, the latter y is in H_{i-1} . Therefore, by construction of l_i 's we have that $x \in H_i$. Thus, by aforementioned remark, we established $H_{i+1} \subset H_i$. We define the sets $B_i; 0 \leq i < \infty$:

$$B_0 = \mathbb{N} \setminus H_1,$$

$$B_1 = H_1 \setminus H_2$$

$$\dots$$

$$B_i = H_i \setminus H_{i+1}$$

$$\dots$$

Clearly we have $B_i \cap B_j = \emptyset, \forall i \neq j$ and $|\mathbb{N} \setminus (\cup_{i=0}^{\infty} B_i)| \leq 1$. The latter is because for every i the second element (in the increasing order) of H_i is $\geq b^i$, therefore if the latter set would contain 2 elements then the second element (in the increasing order) is unbounded.

We define $A_S = \bigcup_{i=0}^{\infty} A_i$, where A_i 's are defined in the following manner:

$$A_0 \doteq S \cap B_0, C_0 \doteq A_0^c \cap B_0$$

$$D_1 \doteq B_1 \setminus \{x \in B_1 \mid ax \in bB_0 + c\}, A_1 \doteq (B_1 \cap \{x \mid ax \in bC_0 + c\}) \cup (D_1 \cap S),$$

$$C_1 \doteq A_1^c \cap B_1$$

...

$$D_i \doteq B_i \setminus \{x \in B_i \mid ax \in bB_{i-1} + c\}, A_i \doteq (B_i \cap \{x \mid ax \in bC_{i-1} + c\}) \cup (D_i \cap S),$$

$$C_i \doteq A_i^c \cap B_i$$

...

Here it is worthwhile to remark that for every $i : A_i \subset B_i$ and $B_i = A_i \cup C_i$. Therefore $A_S \subset \cup_{i=0}^{\infty} B_i$.

If for some i we have $y \in A_i \subset B_i$ then $x : ax = by + c$ satisfies $ax \in bA_i + c$ and by the construction of B_i 's we know that $x \in B_{i+1}$ or $x \in \mathbb{N} \setminus (\cup_{i=0}^{\infty} B_i)$. In the first case $x \notin A_{i+1} \Rightarrow x \notin A_S$. In the second case $x \notin A_S$.

Thus in A_S the equation (3.8) is unsolvable. Our main claim is the following.

For almost every subset S of \mathbb{N} the set A_S is a normal set.

By normality we mean that the infinite binary sequence $1_{A_S} \in \{0, 1\}^{\infty}$ is a normal binary sequence. The probability measure on subsets of \mathbb{N} is the product on $\{0, 1\}^{\infty}$ of probability measures $(\frac{1}{2}, \frac{1}{2})$.

The tool for proving the claim is the following easy lemma (for a proof see appendix, lemma 6.3).

A subset A of natural numbers is a normal set \Leftrightarrow for any $k \in (\mathbb{N} \cup \{0\})$ and any $i_1 < i_2 < \dots < i_k$ we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \chi_A(n) \chi_A(n+i_1) \dots \chi_A(n+i_k) = 0, \quad (3.9)$$

where $\chi_A(n) \doteq 21_A(n) - 1$.

First of all, we denote by $T_N = \frac{1}{N} \sum_{n=1}^N \chi_{A_S}(n) \chi_{A_S}(n+i_1) \dots \chi_{A_S}(n+i_k)$. Because of randomness of S , T_N is a random variable (the probability on Borel subsets of $\{0, 1\}^{\mathbb{N}}$). We will prove that $\sum_{N=1}^{\infty} E(T_N^2) < \infty$ and this will imply that $T_N \rightarrow_{N \rightarrow \infty} 0$ for almost every $S \subset \mathbb{N}$.

$$E(T_N^2) = \frac{1}{N^2} \sum_{n,m=1}^N E(\chi_{A_S}(n) \chi_{A_S}(n+i_1) \dots \chi_{A_S}(n+i_k) \chi_{A_S}(m) \chi_{A_S}(m+i_1) \dots \chi_{A_S}(m+i_k)).$$

A unique possible element of complement of $\cup_i B_i = \cap_{i=1}^{\infty} H_i$ doesn't effect the normality of A_S and we assume without loss of generality that $\cap_{i=1}^{\infty} H_i = \emptyset$, thus $\mathbb{N} = \cup_{i=0}^{\infty} B_i$. For every number $n \in \mathbb{N}$ we define the chain of n , $Ch(n)$, in the following way:

If $n \in B_0$, then $Ch(n) = (n)$.

If $n \in B_1$, then two situations are possible. In the first one there exists a unique $y \in B_0$ such that $an = by + c$. We set $Ch(n) = (n, y) = (n, Ch(y))$. In the second situation we can not find such y from B_0 and we set $Ch(n) = (n)$.

If $n \in B_{i+1}$, then again two situations are possible. In the first one there exists $y \in B_i$ such that $an = by + c$. In this case we set $Ch(n) = (n, Ch(y))$. In the second situation there is no such y from no one of B_0, \dots, B_i . In this case we set $Ch(n) = (n)$. We define the length of $Ch(n)$, $l(n)$, to be a number of elements in $Ch(n)$.

For every $n \in \mathbb{N}$ we define the ancestor of n , $a(n)$, to be the last element of the chain of n ($Ch(n)$). To determine whether or not $n \in A_S$ will depend on whether $a(n) \in S$. The exact relationship will depend on the i for which $n \in B_i$ and j for which $a(n) \in B_j$ or in other words on length of $Ch(n)$: $\chi_{A_S}(n) = (-1)^{i-j} \chi_S(a(n)) = (-1)^{l(n)-1} \chi_S(a(n))$ (as proven below).

We say that n is a descendant of $a(n)$.

We prove the formula $\chi_{A_S}(n) = (-1)^{i-j} \chi_S(a(n))$, where i and j are defined by $n \in B_i$ and $a(n) \in B_j$.

If $i = 0 \Rightarrow j = 0$ and the formula is obvious.

For $i > 0$: If $j = i$ then the formula again is obvious. If $j = i - 1$ then in case $a(n) \in A_{i-1}$ we get that $n \notin A_i$ and in case $a(n) \notin A_{i-1}$ we get that $n \in A_i$. Therefore, we get $\chi_{A_S}(n) = -\chi_S(a(n))$. For general $j < i - 1$ the argument is the same.

It is evident that $E(\chi_{A_S}(n_1) \dots \chi_{A_S}(n_k)) \neq 0 \Leftrightarrow$ every number $a(n_i)$ occurs an even number of times among numbers $a(n_1), a(n_2), \dots, a(n_k)$.

We will bound the number of n, m 's inside the square $[1, N] \times [1, N]$ such that $E(\chi_{A_S}(n) \chi_{A_S}(n + i_1) \dots \chi_{A_S}(n + i_k) \chi_{A_S}(m) \chi_{A_S}(m + i_1) \dots \chi_{A_S}(m + i_k)) \neq 0$.

For a given $n \in [1, N]$ we will count all m 's inside $[1, N]$ such that for the ancestor of n there will be a chance to have a twin among the ancestors of all $n + i_1, \dots, n + i_k, m, m + i_1, \dots, m + i_k$.

First of all it is obvious that in the interval $[1, N]$ for a given ancestor there can be at most $\log_{\frac{b}{a}} N + C_1$ descendants, where C_1 is a constant. For all but a finite number of n 's it is impossible that among $n + i_1, \dots, n + i_k$ there is the same ancestor as for n . Therefore we should focus on ancestors of the set $\{m, m + i_1, \dots, m + i_k\}$. For a given n we might have at most $(k + 1)(\log_{\frac{b}{a}} N + C_1)$ options for the number m to provide that for one of elements of the set $\{m, m + i_1, \dots, m + i_k\}$ has the same ancestor as n . Therefore for most of $n \in [1, N]$ (except maybe a bounded number C_2 of n 's which depends only on $\{i_1, \dots, i_k\}$ and doesn't depend on N) we have at most $(k + 1)(\log_{\frac{b}{a}} N + C_1)$ possibilities for m 's such that

$$E(\chi_{A_S}(n) \chi_{A_S}(n + i_1) \dots \chi_{A_S}(n + i_k) \chi_{A_S}(m) \chi_{A_S}(m + i_1) \dots \chi_{A_S}(m + i_k)) \neq 0.$$

Thus we have

$$E(T_N^2) \leq \frac{1}{N^2} \left(\sum_{n=1}^N (k + 1)(\log_{\frac{b}{a}} N + C_1) + C_2 N \right) = \frac{1}{N} ((k + 1) \log_{\frac{b}{a}} N + C_3),$$

where C_3 is a constant. This means that

$$\sum_{N=1}^{\infty} E(T_{N^2}^2) < \infty.$$

Therefore $T_{N^2} \rightarrow_{N \rightarrow \infty} 0$ for almost every $S \subset \mathbb{N}$. By lemma 6.4 it follows that $T_N \rightarrow_{N \rightarrow \infty} 0$ almost surely.

In the general case, where a, b are not relatively prime, if c satisfies (3.8) then it should be divisible by (a, b) . Therefore by dividing the equation (3.8) by (a, b) we reduce the problem to the previous case.

□

We will use the following notation:

Let W be a subset of \mathbb{Q}^n . Then for any subset $I = \{i_1, \dots, i_p\} \subset \{1, 2, \dots, n\}$ we define

$$Proj_I W = W_I = \{(w_{i_1}, \dots, w_{i_p}) \mid \forall w = (w_1, w_2, \dots, w_n) \in W\}.$$

The next step involves an algebraic statement with a topological proof which we have to establish.

Lemma 3.2.1 *Let W be a non-trivial cone in \mathbb{Q}^n which has the property that for every two vectors $\vec{x}_1 = \{a_1, a_2, \dots, a_n\}^t, \vec{x}_2 = \{b_1, b_2, \dots, b_n\}^t \in W$ there exist two coordinates $1 \leq i < j \leq n$ (depend on the choice of \vec{x}_1, \vec{x}_2) such that*

$$\det \begin{pmatrix} a_i & b_i \\ a_j & b_j \end{pmatrix} = 0,$$

then there exist at least two coordinates $i < j$ such that the projection of W on these two coordinates is of dimension ≤ 1 ($\dim_{\mathbb{Q}} \text{span } Proj_{i,j} W \leq 1$).

Proof. First of all W has a non-empty interior in the topological space $V = \text{Span} W$. Assume that there no exist $i \neq j$ such that the projection of W on these two coordinates is of dimension ≤ 1 . Without loss of generality we assume that $V = \mathbb{Q}^n$. Let fix an arbitrary non-zero element $\vec{x} \in W$. For every $i, j : 1 \leq i < j \leq n$ we define the subspace $U_{i,j} = \{\vec{v} \in V \mid Proj_{i,j} \vec{v} \in \text{Span } Proj_{i,j} \vec{x}\}$.

From assumptions of the lemma it follows that $W = \cup_{i,j: 1 \leq i < j \leq n} (W \cap U_{i,j})$. For every $i \neq j$ we obviously have that the interior of $U_{i,j}$ is empty set. We get a contradiction because a finite union of sets with empty interior can not be equal to a set with non-empty interior.

□

Proof. (of theorem 1.4.1, \Rightarrow)

First of all, we shift the affine space of solutions of equation (3.1) to obtain a vector subspace, denote it U . The linear space U must contain vectors with all positive coordinates. Otherwise, the solution space can have only finitely many positive solutions. Take any WM set and delete a finite number of its elements we obtain a set in which the system is not solvable. But removing a finite number of elements from a WM set does not affect the statistics of the remaining set; therefore, it will be still a WM set. Thus, we can generate a WM set A in which the equation (3.1) is not solvable. The latter contradicts the assumption that the system is solvable within every WM set.

Denote by $W = \{\vec{v} \in U \mid \langle \vec{v}, \vec{e}_i \rangle > 0, \forall i : 1 \leq i \leq k\}$. W is a non-trivial cone. By excluding all coordinates $i : 1 \leq i \leq n$ for which we have $Proj_i W = \{0\}$ we can assume that for every $i : 1 \leq i \leq n$ we have $Proj_i W \neq \{0\}$. By lemma 3.2.1 we deduce that there exist maximal subsets of coordinates F_1, \dots, F_l such that for every $p \in \{1, 2, \dots, l\}$ we have $\forall i, j \in F_p$ the space $V_{i,j} \doteq \text{Span } W_{i,j}$ is one dimensional.

We fix $p : 1 \leq p \leq l$. We should show that the projection on F_p of all solutions of (3.1) is on a shifted diagonal, where a shift is the same for all coordinates in F_p . If the projection of W on coordinates from F_p is not on a diagonal then there exist two coordinates $i < j$ from F_p such that $W_{i,j} = \{(ax, bx) \mid x \in \mathbb{N}\}$ for some $a \neq b$ natural numbers. Therefore the projection of the solutions space of (3.1) on i, j has the form

$\{(ax + f_1, bx + f_2) \mid x \in \mathbb{N}\}$, where f_1, f_2 are integers. From proposition 3.2.1 it follows that for any a, b, c , where $a \neq b$, there exists a WM set A such that the equation $ax = by + c$ is not solvable inside A . This proves the existence of a WM set A such that for every $x \in \mathbb{Z}$ we have $(ax + f_1, bx + f_2) \notin A^2$ (we take a WM set A such that the equation $ax = by + (af_2 - bf_1)$ is unsolvable inside A). To prove that a shift is the same quantity for all coordinates in F_p we merely should know that for any natural number c there exists a WM set A_c such that inside A_c the equation $x - y = c$ is not solvable. The last statement is easy to verify.

Denote by $E = \{1, 2, \dots, k\} \setminus (F_1 \cup F_2 \cup \dots \cup F_l)$. Then there exist two vectors $\vec{x}_1, \vec{x}_2 \in W$ such that the projection of them on two arbitrary coordinates from E is two dimensional. Therefore the condition a) of the theorem 1.4.1 holds. Moreover, by the same argument that was used to extract maximal subsets of coordinates F_1, \dots, F_l and by preceding remarks there exist \vec{x}_1, \vec{x}_2 which satisfy condition a) and additionally satisfy conditions b) and c) of the theorem. This completes the proof.

□

3.3 Comparison with Rado's Theorem

We recall that the problem of solvability of a system of linear equations for any finite partition of \mathbb{N} was solved by Rado in [14]. Such systems of linear equations are called regular (or partition regular). Before citing the theorem we would say that we may expect regular systems to be solvable as well inside every WM set. This is in fact the case and could be shown directly, without use of theorem 1.4.2, by the technique of Furstenberg and Weiss that was developed in their dynamical proof of Rado's theorem (see [11]). Instead of doing so, we obtain this result by use of theorem 1.4.1.

First of all we should describe Rado's regular systems. We will need a definition of the following object.

Definition 3.3.1 *A rational $p \times q$ matrix (a_{ij}) is said to be of level l if the index set $\{1, 2, \dots, q\}$ can be divided into l disjoint subsets I_1, I_2, \dots, I_l and rational numbers c_j^r may be found for $1 \leq r \leq l$ and $1 \leq j \leq q$ such that the following relationships are satisfied:*

$$\begin{aligned} \sum_{j \in I_1} a_{ij} &= 0 \\ \sum_{j \in I_2} a_{ij} &= \sum_{j \in I_1} c_j^1 a_{ij} \\ &\dots \\ \sum_{j \in I_l} a_{ij} &= \sum_{j \in I_1 \cup I_2 \cup \dots \cup I_{l-1}} c_j^{l-1} a_{ij} \end{aligned}$$

for $i = 1, 2, \dots, p$.

Theorem 3.3.1 (Rado) *A system of linear equations is regular if and only if for some l the matrix (a_{ij}) is of level l and it is homogeneous, i.e. a system is of the form*

$$\sum_{j=1}^q a_{ij} x_j = 0, \quad i = 1, 2, \dots, p.$$

After recalling Rado's result we are ready to demonstrate the following.

Proposition 3.3.1 *A regular system is solvable in every WM set.*

Proof. Let a system $\sum_{j=1}^q a_{ij}x_j = 0, i = 1, 2, \dots, p$ be regular. We will use the fact that the system is solvable for any finite partition of \mathbb{N} . First of all, the set of solutions of a regular system is a subspace of \mathbb{Q}^q , let us denote it V . It is obvious that V contains vectors with all positive components. If for some $1 \leq i < j \leq q$ we have $Proj_{i,j}^+ V$ (where $Proj_{i,j}^+ V = \{(x, y) | x, y \geq 0 \ \& \ \exists \vec{v} \in V : \langle \vec{v}, \vec{e}_i \rangle = x, \langle \vec{v}, \vec{e}_j \rangle = y\}$) is contained in a line, then $Proj_{i,j}^+ V$ is diagonal, i.e. is contained in $\{(x, x) | x \in \mathbb{Q}\}$. Otherwise, we can generate a partition of \mathbb{N} into two disjoint sets S_1, S_2 such that no S_1^q and no S_2^q intersects V :

This partition is constructed by an iterative process. Without loss of generality we may assume that the line is $x = ny$, where $n \in \mathbb{N}$. The general case is treated in the same way. We start with $S_1 = S_2 = \emptyset$. Let $1 \in S_1$.

Then we "color" the infinite geometric progression $\{n^m | m \in \mathbb{N}\}$ (adding elements to either S_1 or S_2) in such way that there is no (x, y) on the line from S_1^2, S_2^2 . Then we take a minimal element from \mathbb{N} which is still uncolored. Call it a . Then we add a to S_1 . And again we "color" $\{an^m | m \in \mathbb{N}\}$.

By induction in this way we obtain a desired partition of \mathbb{N} .

This contradicts the assumption that the given system is regular.

If for all $1 \leq i < j \leq k$ we have $\dim_{\mathbb{Q}} \text{span}(Proj_{i,j}^+ V) = 2$ then by lemma 3.2.1 it follows that there exist two vectors $\vec{x}_1, \vec{x}_2 \in V$ which satisfy all requirements of theorem 1.4.1. Thus, in this case the system is solvable in every WM set.

Otherwise, let F_1, \dots, F_l denote maximal subsets of indices such that for every $p \in \{1, \dots, l\}$ we have for every $i \neq j, i, j \in F_p : \dim_{\mathbb{Q}} \text{span}(Proj_{i,j}^+ V) = 1$. Let $E = \{1, 2, \dots, k\} \setminus (F_1 \cup \dots \cup F_l)$. For every $p : 1 \leq p \leq l$ we choose arbitrarily one representative index within F_p and denote it by j_p ($j_p \in F_p$). Then by passing to the subset of indices $I \doteq E \cup \{j_1, \dots, j_l\}$ we can show by use lemma 3.2.1 that there exist $\vec{x}_1, \vec{x}_2 \in V$ with all positive coordinates such that for every $i \neq j, i, j \in I$ we have $\dim_{\mathbb{Q}} Proj_{i,j}(\text{span}(\vec{x}_1, \vec{x}_2)) = 2$. The latter ensures that the vectors \vec{x}_1, \vec{x}_2 satisfy all requirements of theorem 1.4.1 and, therefore, the system is solvable in every WM set.

□

4 An additive analog of polynomial multiple recurrence for WM Sets

We recall the notation which was introduced earlier.

Notation: The Hilbert space $L^2(N)$ is the space of all real-valued functions on the finite set $\{1, 2, \dots, N\}$ endowed with the following scalar product:

$$\langle u, v \rangle_N = \frac{1}{N} \sum_{n=1}^N u(n)v(n).$$

We denote by $\|u\|_N = \sqrt{\langle u, u \rangle_N}$.

The following definition will be used extensively.

Definition 4.0.2 Polynomials $p_1, \dots, p_k \in \mathbb{Z}[n]$ are called **essentially distinct** if the difference of every two of them is non-constant polynomial.

The main result of this chapter is the following

Theorem 1.4.2: For every $k \in \mathbb{N}$ the system

$$\begin{cases} x + y_1 = p_1(z) \\ x + y_2 = p_2(z) \\ \dots \\ x + y_k = p_k(z) \end{cases} \quad (4.1)$$

is solvable within every WM set if $\deg(p_1) = \deg(p_2) = \dots = \deg(p_k)$, p_1, \dots, p_k are essentially distinct and have positive leading coefficients.

4.1 Orthogonality of polynomial shifts

The following lemma is essentially the main tool in the proof of the foregoing theorem. It is inspired by the analogous proposition 3.1.1 in section 3.1.

Lemma 4.1.1 Let $A \subset \mathbb{N}$ be a WM set and assume that $p_1, \dots, p_k \in \mathbb{Z}[n]$ are essentially distinct polynomials with positive leading coefficients. We set $\xi(n) = 1_A(n) - d(A)$ for non-negative n and zero for $n \leq 0$, and we assume $q(n) \in \mathbb{Z}[n]$ with a positive leading coefficient, $\deg(q) \geq \max_{1 \leq i \leq k} \deg(p_i)$ and for every $i : 1 \leq i \leq k$ such that $\deg(p_i) = \deg(q)$ we have that the leading coefficient of $q(n)$ is bigger than that of p_i . Then for every $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for every $J \geq J(\varepsilon)$ there exists $N(J, \varepsilon)$ such that for every $N \geq N(J, \varepsilon)$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^J a_{N+j} \xi(n - p_1(N+j)) \xi(n - p_2(N+j)) \dots \xi(n - p_k(N+j)) \right\|_{q(N)} < \varepsilon$$

for every $\{a_n\} \in \{0, 1\}^{\mathbb{N}}$.

Proof. We prove this statement by using an analog of Bergelson's PET induction, see [2]. Let $F = \{p_1, \dots, p_k\}$ be a finite set of polynomials and assume that the largest of the degrees of p_i equals d . For every $i : 1 \leq i \leq d$ we denote by n_i the number of different groups of polynomials of degree i , where two polynomials p_{j_1}, p_{j_2} of degree i are in the same group if and only if they have the same leading coefficient. We will say that (n_1, \dots, n_d) is the *characteristic vector* of F .

We prove a more general statement than the statement of the lemma.

Let $\mathcal{F}(n_1, \dots, n_d)$ be the family of all finite sets of essentially distinct polynomials having characteristic vector (n_1, \dots, n_d) . Consider the following two statements:

$L(k; n_1, \dots, n_d)$: 'For every $\{g_1, \dots, g_{n_1}, q_1, \dots, q_l\} \in \mathcal{F}(n_1, \dots, n_d)$, where $d \leq \deg(q)$, q is increasing faster than any $q_i, i : 1 \leq i \leq l$ (the exact statement is formulated in lemma) and g_1, \dots, g_{n_1} are linear polynomials, and every $\varepsilon, \delta > 0$ there exists $H(\delta, \varepsilon) \in \mathbb{N}$ such that for every $H \geq H(\delta, \varepsilon)$ there exists $J(H, \varepsilon) \in \mathbb{N}$ such that for every $J \geq J(H, \varepsilon)$ there exists $N(J, H, \varepsilon) \in \mathbb{N}$ such that for every $N \geq N(J, H, \varepsilon)$ for a set of $\{h_1, \dots, h_k\} \in [1 \dots H]^k$ of density at least $1 - \delta$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^J a_{N+j} \prod_{i=1}^{n_1} \prod_{\epsilon \in \{0,1\}^k} \xi(n - g_i(N+j) - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \prod_{i=1}^l \xi(n - q_i(N+j)) \right\|_{q(N)} < \varepsilon,$$

for every $\{a_n\} \in \{0, 1\}^{\mathbb{N}}$.

$L(k; \overline{n_1, \dots, n_i}, n_{i+1}, \dots, n_d)$: ' $L(k; n_1, \dots, n_d)$ is valid for any n_1, \dots, n_i '.

Lemma 4.1.1 is the special case $L(0; \overline{n_1, \dots, n_d})$, where $d \leq \deg(q)$ and the polynomial q is increasing faster than all polynomials in the given family of polynomials which has the characteristic vector (n_1, \dots, n_d) . In order to prove the latter it is enough to establish $L(k; 1), \forall k \in \mathbb{N} \cup \{0\}$, and to prove the following implications:

$$S.1_d : L(k+1; n_1, n_2, \dots, n_d) \Rightarrow L(k; n_1+1, n_2, \dots, n_d); \quad k, n_1, \dots, n_{d-1} \geq 0, n_d \geq 1, d \geq 1$$

$$S.2_{d,i} : L(0; \overline{n_1, \dots, n_{i-1}}, n_i, \dots, n_d) \Rightarrow L(k; \underbrace{0, \dots, 0}_{i-1 \text{ zeros}}, n_i+1, n_{i+1}, \dots, n_d);$$

$$k, n_1, \dots, n_{d-1} \geq 0, n_d \geq 1, d \geq i > 1$$

$$S.3_d : L(k; \overline{n_1, \dots, n_d}) \Rightarrow L(k; \underbrace{0, \dots, 0}_d, 1), \quad k \geq 0, d \geq 1$$

We start with a proof of statement $S.2_{d,i}$. Suppose that F is a finite set of essentially distinct polynomials and assume that the characteristic vector of F equals

$(\underbrace{0, \dots, 0}_{i-1 \text{ zeros}}, n_i+1, n_{i+1}, \dots, n_d)$. Fix any of the n_i+1 groups of polynomials of degree i

and denote its polynomials by g_1, \dots, g_m . Denote the remaining polynomials in F by q_1, \dots, q_l . Because there are no linear polynomials among the polynomials of F , we have to show the following:

Let the family $F \doteq \{g_1, \dots, g_m, q_1, \dots, q_l\}$ of polynomials with the characteristic vector $(\underbrace{0, \dots, 0}_{i-1 \text{ zeros}}, n_i+1, n_{i+1}, \dots, n_d)$, where $\{g_1, g_2, \dots, g_m\} \in \mathbb{Z}[n]$ is one of the groups of F of the degree $i, i > 1$. Let A be a WM set and denote by ξ the normalized WM-sequence,

i.e., $\xi(n) = 1_A(n) - d(A)$, $\forall n \in \mathbb{N}$. For every $\varepsilon, \delta > 0$ there exists $H(\varepsilon, \delta) \in \mathbb{N}$ such that for every $H \geq H(\varepsilon, \delta)$ there exists $J(\varepsilon, H)$ such that for every $J \geq J(\varepsilon, H)$ there exists $N(J, \varepsilon, H)$ such that for every $N \geq N(J, \varepsilon, H)$ for a set of $(h_1, \dots, h_k) \in \{1, \dots, H\}^k$ of density which is at least $1 - \delta$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^J a_{N+j} \prod_{\epsilon \in \{0,1\}^k} \xi(n - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \xi(n - g_1(N+j)) \dots \xi(n - g_m(N+j)) \right\|$$

$$\xi(n - q_1(N+j)) \dots \xi(n - q_l(N+j)) \Big\|_{q(N)} < \varepsilon,$$

for every $\{a_n\} \in \{0,1\}^{\mathbb{N}}$ and with the condition $\deg(q) \geq d$ and q is increasing faster than any q_i , $i : 1 \leq i \leq l$.

Denote by

$$u_j(n) \doteq a_{N+j} \xi(n - g_1(N+j)) \dots \xi(n - g_m(N+j))$$

$$\xi(n - q_1(N+j)) \dots \xi(n - q_l(N+j)),$$

$$w(n) = \prod_{\epsilon \in \{0,1\}^k} \xi(n - \epsilon_1 h_1 - \dots - \epsilon_k h_k),$$

$$v_j(n) = w(n) u_j(n),$$

$$n = 1, \dots, q(N).$$

The sequence $w(n)$ is bounded by 1 and therefore to prove that $\left\| \frac{1}{J} \sum_{j=1}^J v_j \right\|_{q(N)}$ is small it is sufficient to show that $\left\| \frac{1}{J} \sum_{j=1}^J u_j \right\|_{q(N)}$ is small.

We apply the van der Corput lemma (see lemma 6.1 in appendix):

$$\begin{aligned} & \frac{1}{J} \sum_{j=1}^J \langle u_j, u_{j+h} \rangle_{q(N)} = \\ & \frac{1}{q(N)} \sum_{n=1}^{q(N)} \frac{1}{J} \sum_{j=1}^J a_{N+j} \xi(n - g_1(N+j)) \dots \xi(n - g_m(N+j)) \\ & \quad \xi(n - q_1(N+j)) \dots \xi(n - q_l(N+j)) \\ & \quad a_{N+j+h} \xi(n - g_1(N+j+h)) \dots \xi(n - g_m(N+j+h)) \\ & \quad \xi(n - q_1(N+j+h)) \dots \xi(n - q_l(N+j+h)) = \\ & \frac{1}{q(N) - g_1(N)} \sum_{n=1}^{q(N)} \xi(n) \frac{1}{J} \sum_{j=1}^J a_{N+j} a_{N+j+h} \xi(n - (g_2(N+j) - g_1(N+j))) \dots \\ & \quad \xi(n - (g_m(N+j) - g_1(N+j))) \xi(n - (q_1(N+j) - g_1(N+j))) \dots \\ & \quad \xi(n - (q_l(N+j) - g_1(N+j))) \xi(n - (g_1(N+j+h) - g_1(N+j))) \dots \\ & \quad \xi(n - (g_m(N+j+h) - g_1(N+j))) \xi(n - (q_1(N+j+h) - g_1(N+j))) \dots \\ & \quad \xi(n - (q_l(N+j+h) - g_1(N+j))) + \delta_{N,J} = \\ & \frac{1}{q(N)} \sum_{n=1}^{q(N) - g_1(N)} \xi(n) \frac{1}{J} \sum_{j=1}^J b_{N+j} \xi(n - r_1(N+j)) \dots \xi(n - r_{m-1}(N+j)) \xi(n - r_m(N+j)) \dots \end{aligned}$$

$$\begin{aligned} & \xi(n - r_{m+l-1}(N + j))\xi(n - r_{m+l}(N + j)) \dots \xi(n - r_{2m+l-1}(N + j)) \\ & \xi(n - r_{2m+l}(N + j)) \dots \xi(n - r_{2m+2l-1}(N + j)) + \delta_{N,J}, \end{aligned}$$

where in the second equality we used a change of variable $n \leftarrow n = n - g_1(N + j)$, $b_{N+j} = a_{N+j}a_{N+j+h}$, $\delta_{N,J} \rightarrow \frac{J}{N} \rightarrow 0$ and

$$\begin{cases} r_t(n) = g_{t+1}(n) - g_1(n), & t : 1 \leq t \leq m-1 \\ r_t(n) = q_{t-(m-1)}(n) - g_1(n), & t : m \leq t \leq m+l-1 \\ r_t(n) = g_{t-(m+l-1)}(n+h) - g_1(n), & t : m+l \leq t \leq 2m+l-1 \\ r_t(n) = q_{t-(2m+l-1)}(n+h) - g_1(n), & t : 2m+l \leq t \leq 2m+2l-1. \end{cases}$$

For all but a finite number of h 's the polynomials $\{r_t(n)\}_{t=1}^{2m+2l-1}$ are essentially distinct, because $i > 1$ and the polynomials $g_1, \dots, g_m, q_1, q_l$ are essentially distinct. To see the last property we notice that if we take two polynomials r_t 's from the same group (there are 4 groups), then their difference is a non-constant because the initial polynomials are essentially distinct. If we take two polynomials from different groups then three cases are possible. In the first case the difference of these polynomials is $g_t(n+h) - g_t(n)$ or $q_t(n+h) - q_t(n)$ for some t . We assume that $i > 1$ therefore $\min_{1 \leq t \leq l} \min(\deg(q_t), \deg(g_1)) > 1$ and from this it follows that $g_t(n+h) - g_t(n)$ and $q_t(n+h) - q_t(n)$ are non-constant polynomials. In the second case we get for some $t_1 \neq t_2$: $g_{t_1}(n+h) - g_{t_2}(n)$ or $q_{t_1}(n+h) - q_{t_2}(n)$. Here we note that the map $h \mapsto p(n+h)$ is an injective map from \mathbb{N} to the set of essentially distinct polynomials, if $\deg(p) > 1$. Thus, for all but a finite number of h 's we get again a non-constant difference. In the third case we get for some t_1, t_2 : $g_{t_1}(n+h) - q_{t_2}(n)$ or $q_{t_1}(n+h) - g_{t_2}(n)$. The resulting polynomial has the same degree as q_t .

The characteristic vector of the set of polynomials $\{r_1, \dots, r_{2m+2l-1}\}$ has the form $(c_1, \dots, c_{i-1}, n_i, n_{i+1}, \dots, n_d)$. The polynomials from the second and the fourth group have the same degree as q_t and the same leading coefficient as q_t if $\deg(q_t) > \deg(g_1)$ and the leading coefficient will be the difference of leading coefficients of q_t and g_1 if $\deg(q_t) = \deg(g_1)$. The polynomials from the first and the third group will be of degree smaller than $\deg(g_1)$.

Applying $L(0; \overline{n_1}, \dots, \overline{n_{i-1}}, n_i, \dots, n_d)$ with the new polynomial $q(n) - g_1(n)$ which is increasing faster than all the polynomials $\{r_t(n)\}_{t=1}^{2m+2l-1}$ and the Cauchy-Schwartz inequality we get that for all but a finite number of h 's and for every $\varepsilon > 0$ there exists $J(\varepsilon, h)$ such that for every $J \geq J(\varepsilon, h)$ there exists $N(J, \varepsilon, h)$ such that for every $N \geq N(J, \varepsilon, h)$ we have

$$\left| \frac{1}{J} \sum_{j=1}^J < u_j, u_{j+h} >_{q(N)} \right| < \varepsilon,$$

for every $\{a_n\} \in \{0, 1\}^{\mathbb{N}}$.

By the van der Corput lemma it follows that for every $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for every $J \geq J(\varepsilon)$ there exists $N(J, \varepsilon)$ such that for every $N \geq N(J, \varepsilon)$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^J u_j \right\|_{q(N)} < \varepsilon,$$

for every $\{a_n\} \in \{0,1\}^{\mathbb{N}}$. Thus we have shown the validity of $L(k; \underbrace{0, \dots, 0}_{i-1 \text{ zeros}}, n_i + 1, n_{i+1}, \dots, n_d)$.

We proceed with a proof of $S.1_d$. We fix the $n_1 + 1$ groups of the polynomials of degree 1 and denote its polynomials by $g_1(n) = c_1n + d_1, \dots, g_{n_1+1} = c_{n_1+1}n + d_1$. (By the assumption that all given polynomials are essentially distinct we get that in any group of degree 1 there is only one polynomial). The remaining polynomials we denote by q_1, \dots, q_l . The set of polynomials $\{g_1, \dots, g_{n_1+1}, q_1, \dots, q_l\}$ has the characteristic vector $(n_1 + 1, n_2, \dots, n_d)$. Again we apply the van der Corput lemma. Let $u_j(n)$ be defined as following

$$u_j(n) \doteq a_{N+j} \prod_{i=1}^{n_1+1} \prod_{\epsilon \in \{0,1\}^k} \xi(n - g_i(N+j) - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \prod_{i=1}^l \xi(n - q_i(N+j)),$$

$$n = 1, \dots, q(N).$$

Then we have

$$\begin{aligned} & \frac{1}{J} \sum_{j=1}^J < u_j, u_{j+h} >_{q(N)} = \\ & \frac{1}{q(N)} \sum_{n=1}^{q(N)} \frac{1}{J} \sum_{j=1}^J a_{N+j} a_{N+j+h} \\ & \prod_{i=1}^{n_1+1} \prod_{\epsilon \in \{0,1\}^k} \xi(n - g_i(N+j) - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \prod_{i=1}^l \xi(n - q_i(N+j)) \\ & \prod_{i=1}^{n_1+1} \prod_{\epsilon \in \{0,1\}^k} \xi(n - g_i(N+j+h) - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \prod_{i=1}^l \xi(n - q_i(N+j+h)) = \\ & \frac{1}{q(N) - g_1(N)} \sum_{n=1}^{q(N)} \prod_{\epsilon \in \{0,1\}^k} \xi(n - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \xi(n - \epsilon_1 h_1 - \dots - \epsilon_k h_k - c_1 h) \\ & \frac{1}{J} \sum_{j=1}^J b_{N+j} \prod_{i=1}^{n_1} \prod_{\epsilon \in \{0,1\}^k} \xi(n - (c_{i+1} - c_1)(N+j) - (d_{i+1} - d_1) - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \\ & \prod_{i=1}^{n_1} \prod_{\epsilon \in \{0,1\}^k} \xi(n - (c_{i+1} - c_1)(N+j) - (d_{i+1} - d_1) - \epsilon_1 h_1 - \dots - \epsilon_k h_k - c_{i+1} h) \\ & \prod_{i=1}^l \xi(n - (q_i(N+j) - g_1(N+j))) \prod_{i=1}^l \xi(n - (q_i(N+j+h) - g_1(N+j))) + \delta_{N,J}, \end{aligned}$$

where in the second equality we made a change of variable $n \leftarrow n - g_1(N+j)$ and $b_{N+j} = a_{N+j} a_{N+j+h}$, $\delta_{N,J} \rightarrow \frac{J}{N} \rightarrow 0$.

Denote by $r_i(n) = (c_{i+1} - c_1)n + (d_{i+1} - d_1)$, $i : 1 \leq i \leq n_1$, $s_i(n) = q_i(n) - g_1(n)$, $t_i(n) = q_i(n + h) - g_1(n)$, $i : 1 \leq i \leq l$. Then the last expression may be rewritten as

$$\begin{aligned} & \frac{1}{q(N)} \sum_{n=1}^{q(N)-g_1(N)} \prod_{\epsilon \in \{0,1\}^k} \xi(n - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \xi(n - \epsilon_1 h_1 - \dots - \epsilon_k h_k - c_1 h) \\ & \frac{1}{J} \sum_{j=1}^J b_{N+j} \prod_{i=1}^{n_1} \prod_{\epsilon \in \{0,1\}^k} \xi(n - r_i(N+j) - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \\ & \quad \xi(n - r_i(N+j) - \epsilon_1 h_1 - \dots - \epsilon_k h_k - c_{i+1} h) \\ & \prod_{i=1}^l \xi(n - s_i(N+j)) \xi(n - t_i(N+j)) + \delta_{N,J} \doteq E1 + \delta_{N,J}. \end{aligned}$$

For every $i : 1 \leq i \leq l$ the polynomials s_i, t_i are in the same group (have the same degree and the same leading coefficient), therefore the characteristic vector of the family $\{s_1, t_1, \dots, s_l, t_l\}$ is the same as of the family $\{s_1, s_2, \dots, s_l\}$ and, obviously, the characteristic vector of the latter family is the same as of the family $\{q_1, q_2, \dots, q_l\}$ and is equal to $(0, n_2, n_3, \dots, n_d)$. Again the polynomial $q(n) - g_1(n)$ is increasing faster than any polynomial in the family $\{s_1, t_1, \dots, s_l, t_l\}$. By use of $L(k+1; n_1, \dots, n_d)$ and the Cauchy-Schwartz inequality we show that $|E1|$ is arbitrarily small for a set of arbitrarily large density of (h_1, \dots, h_k, h) 's. Therefore, by the van der Corput lemma we deduce the validity of $L(k; n_1 + 1, n_2, \dots, n_d)$.

The proof of $S.3_d$ goes exactly in the same way as that of $S.2_{d,i}$.

Proof of $L(k; 1)$, $\forall k \in \mathbb{N} \cup 0$:

Assume that $g_1(n) = c_1 n + d_1$, $c_1 > 0$ and $\deg(q) > 1$. We show that

For every $\varepsilon, \delta > 0$ there exists $H(\delta, \varepsilon) \in \mathbb{N}$ such that for every $H \geq H(\delta, \varepsilon)$ there exists $J(H, \varepsilon) \in \mathbb{N}$ such that for every $J \geq J(H, \varepsilon)$ there exists $N(J, H, \varepsilon)$ such that for every $N \geq N(J, H, \varepsilon)$ we have for a set of $(h_1, \dots, h_k) \in \{1, \dots, H\}^k$ of density which is at least $1 - \delta$ the following

$$\left\| \frac{1}{J} \sum_{j=1}^J a_{N+j} \prod_{\epsilon \in \{0,1\}^k} \xi(n - g_1(N+j) - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \right\|_{q(N)} < \varepsilon$$

for every $\{a_n\} \in \{0, 1\}^{\mathbb{N}}$.

We recall that to a WM set A is associated the weakly-mixing system $(X_\xi, \mathbb{B}, T, \mu)$. We define the function f on X_ξ by the following rule: $f(\omega) = \omega_0$, $\omega = \{\omega_0, \dots, \omega_n, \dots\} \in X_\xi$. It is evident that f is continuous and $\int_{X_\xi} f(x) d\mu(x) = 0$. By genericity of the point $\xi \in X_\xi$ we get

$$\frac{q(N)}{q(N) - g_1(N)} \left\| \frac{1}{J} \sum_{j=1}^J a_{N+j} \prod_{\epsilon \in \{0,1\}^k} \xi(n - g_1(N+j) - \epsilon_1 h_1 - \dots - \epsilon_k h_k) \right\|_{q(N)}^2 \rightarrow_{N \rightarrow \infty} 0$$

$$\int_{X_\xi} \left(\frac{1}{J} \sum_{j=1}^J a_{N+J+1-j} T^{c_1 j} \left(\prod_{\epsilon \in \{0,1\}^k} T^{\epsilon_1 h_1 + \dots + \epsilon_k h_k} f(x) \right) \right)^2 d\mu(x). \quad (4.2)$$

Denote by g_{h_1, \dots, h_k} the following function on X_ξ :

$$g_{h_1, \dots, h_k}(x) = \prod_{\epsilon \in \{0,1\}^k} T^{\epsilon_1 h_1 + \dots + \epsilon_k h_k} f(x), \quad \forall x \in X_\xi.$$

Then we use the following statement which can be viewed as a corollary of theorem 13.1 of Host and Kra in [12] ($\int_{X_\xi} f(x) d\mu(x) = 0$).

For every $\varepsilon, \delta > 0$ there exists $H(\delta, \varepsilon) \in \mathbb{N}$ such that for every $H \geq H(\delta, \varepsilon)$ for a set of $(h_1, \dots, h_k) \in \{1, \dots, H\}^k$ which has density at least $1 - \delta$ we have

$$\left| \int_{X_\xi} g_{h_1, \dots, h_k}(x) d\mu(x) \right| < \varepsilon.$$

Let $\varepsilon, \delta > 0$. By the foregoing statement there exists $H(\delta, \varepsilon) \in \mathbb{N}$ such that for every $H \geq H(\delta, \varepsilon)$ the set of those $(h_1, \dots, h_k) \in \{1, \dots, H\}^k$ such that

$$\left| \int_{X_\xi} g_{h_1, \dots, h_k}(x) d\mu(x) \right| < \frac{\varepsilon}{4}$$

has density at least $1 - \delta$.

Lemma 6.2 implies that there exists $J(H, \varepsilon) \in \mathbb{N}$ such that for every $J \geq J(H, \varepsilon)$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^J b_j T^{c_1 j} \left(g_{h_1, \dots, h_k}(x) - \int_{X_\xi} g_{h_1, \dots, h_k}(x) d\mu(x) \right) \right\|_{L^2(X_\xi)} < \frac{\varepsilon}{4}$$

for any sequence $\{b_n\} \in \{0, 1\}^\mathbb{N}$.

Therefore, by merging the two last statements we conclude that there exists $H(\delta, \varepsilon) \in \mathbb{N}$ such that for every $H \geq H(\delta, \varepsilon)$ there exists $J(H, \varepsilon) \in \mathbb{N}$ such that for every $J \geq J(H, \varepsilon)$ and for a set of $(h_1, \dots, h_k) \in \{1, \dots, H\}^k$ which has density at least $1 - \delta$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^J b_j T^{c_1 j} g_{h_1, \dots, h_k}(x) \right\|_{L^2(X_\xi)} < \frac{\varepsilon}{2}$$

for any sequence $\{b_n\} \in \{0, 1\}^\mathbb{N}$.

Finally, by use of (4.2), the fact that $\lim_{N \rightarrow \infty} \frac{q(N)}{q(N) - g_1(N)} > 0$ and the last statement we deduce the validity of $L(k; 1)$.

□

The next lemma is a simple consequence of the previous one and is used in the next subsection to prove theorem 1.4.2.

Lemma 4.1.2 Let $A \subset \mathbb{N}$ be a WM set and $p_1, \dots, p_k \in \mathbb{Z}[n]$ are essentially distinct polynomials of the same degree d greater than 1, with positive leading coefficients such that $p_1(n) > p_i(n)$, $\forall 1 < i \leq k$ for sufficiently large n . Then for every $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for every $J \geq J(\varepsilon)$ there exists $N(J, \varepsilon)$ such that for every $N \geq N(J, \varepsilon)$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^J a_{N+j} \xi(p_1(N+j) - n) \xi(p_2(N+j) - n) \dots \xi(p_k(N+j) - n) \right\|_{p_1(N)} < \varepsilon$$

for every $\{a_n\} \in \{0, 1\}^{\mathbb{N}}$, where $\xi(n) = 1_A(n) - d(A)$ for non-negative n 's and zero for $n \leq 0$.

Remark 4.1.1 The lemma is true for the linear case as well, but a proof demands additional technical efforts.

Proof. For a family of polynomials $F = \{p_1, \dots, p_k\}$ with a maximal degree d denote by n_d the number of different leading coefficients of polynomials of degree d from the family F .

As in the proof of lemma 4.1.1 we fix one of the groups of polynomials of degree d (all polynomials in the same group have the same leading coefficient). Assume that the group $\{g_1, \dots, g_m\}$ has the maximal leading coefficient among all polynomials p_1, \dots, p_k . The rest of the polynomials we denote by q_1, \dots, q_l . Without loss of generality assume that $p_1 = g_1, \dots, p_m = g_m$. Denote by $u_j(n)$, $1 \leq n \leq p_1(N)$ the following expression

$$u_j(n) = a_{N+j} \xi(p_1(N+j) - n) \xi(p_2(N+j) - n) \dots \xi(p_k(N+j) - n).$$

For u_j 's we get

$$\begin{aligned} \frac{1}{J} \sum_{j=1}^J \langle u_j, u_{j+h} \rangle_{p_1(N)} &= \frac{1}{p_1(N)} \sum_{n=1}^{p_1(N)} \frac{1}{J} \sum_{j=1}^J a_{N+j} \xi(p_1(N+j) - n) \dots \\ &\quad \xi(p_k(N+j) - n) a_{N+j+h} \xi(p_1(N+j+h) - n) \dots \xi(p_k(N+j+h) - n) = \\ &\quad \frac{1}{p_1(N)} \sum_{n=1}^{p_1(N)} \xi(n) \frac{1}{J} \sum_{j=1}^J b_{N+j} \prod_{i=1}^{m-1} \xi(n - (p_1(N+j) - p_{i+1}(N+j))) \\ &\quad \prod_{i=1}^l \xi(n - (p_1(N+j) - q_i(N+j))) \prod_{i=1}^m \xi(n - (p_1(N+j) - p_i(N+j+h))) \\ &\quad \prod_{i=1}^l \xi(n - (p_1(N+j) - q_i(N+j+h))) + \delta_{J,N}, \end{aligned}$$

where $b_n = a_n a_{n+h}$ and $\delta_{J,N} \rightarrow \frac{J}{N} \rightarrow 0$.

Denote by $r_i(n) = p_1(n) - q_i(n)$; $s_i(n) = p_1(n) - q_i(n+h)$, $i : 1 \leq i \leq l$ and $t_i(n) = p_1(n) - p_i(n)$; $f_i(n) = p_1(n) - p_i(n+h)$, $i : 1 \leq i \leq m$. Then for all but

a finite number of h 's the polynomials $\tilde{F} \doteq \{r_1, \dots, r_l, s_1, \dots, s_l, t_2, \dots, t_m, f_1, \dots, f_m\}$ are essentially distinct and among them the polynomials of degree d have n_d different leading coefficients. Therefore by lemma 4.1.1 for all but a finite number of h 's the following expression is as small as we wish for appropriately chosen J, N .

$$\left\| \frac{1}{J} \sum_{j=1}^J b_{N+j} \prod_{i=1}^{m-1} \xi(n - t_{i+1}(N+j)) \prod_{i=1}^l \xi(n - r_i(N+j)) \right. \\ \left. \prod_{i=1}^m \xi(n - f_i(N+j)) \prod_{i=1}^l \xi(n - s_i(N+j)) \right\|_{p_1(N)}.$$

Finally by Cauchy-Schwartz inequality and van der Corput's lemma we get the desired conclusion. \square

4.2 Proof of theorem 1.4.2

Proof. (of theorem 1.4.2)

The linear case (the degree of the polynomials is 1) follows from theorem 1.4.1:

Denote $p_i(z) = c_i z + d_i$, $\forall i : 1 \leq i \leq k$. We choose the following order of the variables (x, z, y_1, \dots, y_k) . The affine space of the solutions of the additive system (4.1) is $\{(x, z, x - c_1 z - d_1, \dots, x - c_k z - d_k) \mid x, z \in \mathbb{Q}\}$. If we take the vectors $\vec{x}_1 = (1, 0, 1, 1, \dots, 1)$, $\vec{x}_2 = (0, 1, c_1, c_2, \dots, c_k)$, $\vec{f} = (0, 0, -d_1, -d_2, \dots, -d_k)$ then all the requirements of theorem 1.4.1 regarding the system (4.1) are valid. Thus the system (4.1) is solvable within every WM set.

Assume we have an arbitrary WM set A and k essentially distinct polynomials $p_1, \dots, p_k \in \mathbb{Z}[n]$ (a difference of any two of them is a non constant polynomial) of the same degree $d > 1$ with positive leading coefficients and assume that for sufficiently large n 's we have $p_1(n) > p_i(n)$, $\forall i : 2 \leq i \leq k$. Let's define the set F of all z 's where the statement of the theorem fails, namely,

$$F \doteq \{z \in \mathbb{N} \mid \text{for any } (x, y_1, \dots, y_k) \in A^{k+1} \text{ the system 4.1 fails to hold}\}.$$

We shall prove that $d^*(F) = 0$. Since $d(A) > 0$ we can find $z \in A, z \notin F$ and this will yield a solution to (4.1).

Denote by $\{a_n\}$ the indicator sequence of F , i.e., $a_n = 1_F(n)$. We define the sequence ξ to be a normalized indicator sequence of A : $\xi(n) = 1_A(n) - d(A)$, $n \in \mathbb{N}$ and zero for non-positive values of n , where $d(A)$ is the density of A which exists.

We define the expression $B_{N,J}$ to be

$$B_{N,J} \doteq \frac{1}{p_1(N)} \sum_{n=1}^{p_1(N)} \frac{1}{J} \sum_{j=1}^J a_{N+j} 1_A(n) 1_A(p_1(N+j) - n) \quad (4.3) \\ 1_A(p_2(N+j) - n) \dots 1_A(p_{k-1}(N+j) - n) \xi(p_k(N+j) - n).$$

Suppose that we have $d^*(F) > 0$. Then there exist intervals $I_{l,J} = [u_{l,J} + 1, u_{l,J} + J]$ (for J big enough) such that $u_{l,J} \rightarrow_{l \rightarrow \infty} \infty$ and $\frac{|F \cap I_{l,J}|}{J} > \frac{d^*(F)}{2}$ for every l and J big enough. By induction on k and i we prove the validity of the following claim.

Claim 1: For every $i : 0 \leq i \leq k - 1$ and every $\varepsilon > 0$ there exist J, l big enough such that

$$\left| \frac{1}{p_1(u_{l,J})} \sum_{n=1}^{p_1(u_{l,J})} \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(n) 1_A(p_1(u_{l,J}+j)-n) \dots \right.$$

$$\left. 1_A(p_i(u_{l,J}+j)-n) \xi(p_{i+1}(u_{l,J}+j)-n) \dots \xi(p_k(u_{l,J}+j)-n) \right| < \varepsilon$$

for every $\{0, 1\}$ -valued sequence $\{b_n\}$.

A proof of claim 1 is by induction on i and k .

In the sequel we use the notation $\langle 1_A, f(n) \rangle_N$, where $f(n)$ is defined for all $n = 1, 2, \dots, N$; which has the same meaning as $\langle 1_A, f \rangle_N = \frac{1}{N} \sum_{n=1}^N 1_A(n) f(n)$.

For $i = 0$ and every k the statement is exactly of lemma 4.1.2. For every $i < k - 1$ we will prove the statement of the claim for $i + 1$ and k provided the statement for i and k , and for $i, k - 1$:

$$\left| \frac{1}{p_1(u_{l,J})} \sum_{n=1}^{p_1(u_{l,J})} \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(n) 1_A(p_1(u_{l,J}+j)-n) \dots \right.$$

$$\left. 1_A(p_i(u_{l,J}+j)-n) 1_A(p_{i+1}(u_{l,J}+j)-n) \xi(p_{i+2}(u_{l,J}+j)-n) \dots \xi(p_k(u_{l,J}+j)-n) \right| =$$

$$\left| < 1_A, \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(p_1(u_{l,J}+j)-n) \dots \right.$$

$$\left. 1_A(p_i(u_{l,J}+j)-n) (\xi(p_{i+1}(u_{l,J}+j)-n) + d(A)) \xi(p_{i+2}(u_{l,J}+j)-n) \dots \right.$$

$$\left. \xi(p_k(u_{l,J}+j)-n) \right\rangle_{p_1(u_{l,J})} \left| \leq \right.$$

$$\left| < 1_A, \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(p_1(u_{l,J}+j)-n) \dots \right.$$

$$\left. 1_A(p_i(u_{l,J}+j)-n) \xi(p_{i+1}(u_{l,J}+j)-n) \xi(p_{i+2}(u_{l,J}+j)-n) \dots \right.$$

$$\left. \xi(p_k(u_{l,J}+j)-n) \right\rangle_{p_1(u_{l,J})} \left| + \right.$$

$$\left. d(A) \right| < 1_A, \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(p_1(u_{l,J}+j)-n) \dots$$

$$\left. 1_A(p_i(u_{l,J}+j)-n) \xi(p_{i+2}(u_{l,J}+j)-n) \dots \right.$$

$$\left. \xi(p_k(u_{l,J}+j)-n) \right\rangle_{p_1(u_{l,J})} \left| < \varepsilon, \right.$$

for big enough J, l . The first summand is small by the statement of the claim for i and k , and the second summand is small by the statement of the claim for i and $k - 1$. This ends the proof of claim 1.

We will use the statement of claim 1 for $i = k - 1$ and we call the statement claim 2.

Claim 2: For every $\varepsilon > 0$ there exist J, l big enough such that the expression

$$\left| \frac{1}{p_1(u_{l,J})} \sum_{n=1}^{p_1(u_{l,J})} \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(n) 1_A(p_1(u_{l,J}+j)-n) \dots \right.$$

$$|1_A(p_{k-1}(u_{l,J} + j) - n)\xi(p_k(u_{l,J} + j) - n)| < \varepsilon$$

for every $\{0,1\}$ -valued sequence $\{b_n\}$.

The next statement enables us to conclude about a boundedness away from zero of $B_{u_{l,J},J}$.

Claim 3: For every $\delta > 0$ for big enough J, l the expression

$$\frac{1}{p_1(u_{l,J})} \sum_{n=1}^{p_1(u_{l,J})} \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(n) 1_A(p_1(u_{l,J} + j) - n) \dots 1_A(p_k(u_{l,J} + j) - n)$$

is bigger than $c(1 - \delta)d^{k+1}(A)\frac{d^*(F)}{3}$, where $c = \min_{2 \leq i \leq k-1} \frac{c_i}{c_1}$ (c_i is a leading coefficient of polynomial p_i) for every $\{0,1\}$ -valued sequence $\{b_n\}$ which has density bigger than $\frac{d^*(F)}{2}$ on all intervals $I_{l,J}$.

The proof is by induction on k .

For $k = 1$ then by using lemma 4.1.2 we have that for J and l big enough

$$\begin{aligned} & \frac{1}{p_1(u_{l,J})} \sum_{n=1}^{p_1(u_{l,J})} \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(n) 1_A(p_1(u_{l,J} + j) - n) = \\ & < 1_A, \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} (\xi(p_1(u_{l,J} + j) - n) + d(A)) >_{p_1(u_{l,J})} \geq \\ & -\varepsilon + d(A) < 1_A, \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} >_{p_1(u_{l,J})} > (1 - \delta)d(A)^2 \frac{d^*(F)}{3}. \end{aligned}$$

Assume the statement of the claim holds for k . Let $(p_1, \dots, p_k, p_{k+1})$ be polynomials of the same degree such that p_1 is the "biggest" among them (see conditions of lemma 4.1.2). Without loss of generality we can assume that $\min_{2 \leq i \leq k+1} c_i = c_{k+1}$. Then for sufficiently large J and l

$$\begin{aligned} & \frac{1}{p_1(u_{l,J})} \sum_{n=1}^{p_1(u_{l,J})} \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(n) 1_A(p_1(u_{l,J} + j) - n) \dots \\ & 1_A(p_k(u_{l,J} + j) - n) 1_A(p_{k+1}(u_{l,J} + j) - n) = \\ & < 1_A, \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(p_1(u_{l,J} + j) - n) \dots \\ & 1_A(p_k(u_{l,J} + j) - n) (\xi(p_{k+1}(u_{l,J} + j) - n) + d(A)) >_{p_1(u_{l,J})} - \\ & d(A) \frac{1}{p_1(u_{l,J})} \sum_{n=p_{k+1}(u_{l,J})}^{p_1(u_{l,J})} 1_A(n) \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(p_1(u_{l,J} + j) - n) \dots 1_A(p_k(u_{l,J} + j) - n) = \\ & d(A) < 1_A, \frac{1}{J} \sum_{j=1}^J b_{u_{l,J}+j} 1_A(p_1(u_{l,J} + j) - n) \dots 1_A(p_k(u_{l,J} + j) - n) >_{p_1(u_{l,J})} + \end{aligned}$$

$$\begin{aligned}
& < 1_A, \frac{1}{J} \sum_{j=1}^J b_{u_l, J+j} 1_A(p_1(u_l, J+j) - n) \dots 1_A(p_k(u_l, J+j) - n) \xi(p_{k+1}(u_l, J+j) - n) >_{p_1(u_l, J)} - \\
& d(A) \frac{1}{p_1(u_l, J)} \sum_{n=p_{k+1}(u_l, J)}^{p_1(u_l, J)} 1_A(n) \frac{1}{J} \sum_{j=1}^J b_{u_l, J+j} 1_A(p_1(u_l, J+j) - n) \dots 1_A(p_k(u_l, J+j) - n) \\
& > \\
& d(A) \frac{1}{p_1(u_l, J)} \sum_{n=1}^{p_{k+1}(u_l, J)} 1_A(n) \frac{1}{J} \sum_{j=1}^J b_{u_l, J+j} 1_A(p_1(u_l, J+j) - n) \dots 1_A(p_k(u_l, J+j) - n) - \varepsilon \\
& > d(A) c(1 - \delta') d(A)^{k+1} \frac{d^*(F)}{3} \\
& > c(1 - \delta) d(A)^{k+2} \frac{d^*(F)}{3}.
\end{aligned}$$

We used claim 2 in the first inequality and induction hypothesis in the second inequality. This ends the proof of claim 3.

By the definition of F it follows that for every non-zero value of

$$a_{u_l, J+j} 1_A(n) 1_A(p_1(u_l, J+j) - n) 1_A(p_2(u_l, J+j) - n) \dots 1_A(p_{k-1}(u_l, J+j) - n)$$

(thus it equals to one), the remaining factor in the summands of $B_{u_l, J, J}$ is negative, namely, $\xi(p_k(u_l, J+j) - n) = -d(A)$. Therefore, by using claim 3 we get $|B_{u_l, J, J}| \geq c(1 - \varepsilon) d^{k+1}(A) \frac{d^*(F)}{3}$ for any l and for J big enough. Thus $|B_{u_l, J, J}|$ is bounded from zero.

On the other hand, by claim 2 it follows that for any $\varepsilon > 0$ there exists $J = J(\varepsilon)$ and $N = N(J(\varepsilon))$ such that $|B_{N, J}| < \varepsilon$. Therefore we get a contradiction.

We have proved that the set of all z 's such that the statement of the theorem holds has a lower density one. Therefore it intersects every set of positive density (even of positive upper density), in particular, A .

□

5 The equation $xy = z$ and normal sets

5.1 Normal sets and diophantine equations

We expect that there are many diophantine equations which are solvable in every normal set. We denote by DSN the family of diophantine equations (including systems of equations) which are solvable within every normal set.

It is easily seen that the equation $x + y = z$ is in DSN . This equation is called the additive Schur equation. Schur proved that the equation is "partition regular". This means that for any finite coloring of \mathbb{N} , there exists a monochromatic solution for the equation (see [15]).

Systems of linear diophantine equations that are partition regular are classified by Rado in [14]. Such systems are usually called Rado systems.

In section 3.3 we show that any Rado system of linear equations is in DSN .

At the moment, we don't know the richness of the DSN family. By the aforementioned result, a large family of linear equations (Rado's systems) are in DSN . For non-linear case, we don't know much. For instance, it is not known whether the equation $x^2 + y^2 = z^2$ is in DSN . In this chapter we prove that the equation $xy = z$ is not in DSN . The last equation is called the multiplicative Schur equation. It should be mentioned that for partitions of \mathbb{N} into finite number of subsets, at least one of subsets contains solutions for both Schur's additive and multiplicative equations (see [4]). Therefore, there exist partition regular equations that are not in DSN . We use the notion of Liouville's function to construct a normal set in which the multiplicative Schur's equation is unsolvable.

Definition 5.1.1 *Liouville's function $\lambda : \mathbb{N} \rightarrow \{-1, 1\}$ is defined as follows:*

$$\lambda(p_1^{e_1} p_2^{e_2} \dots p_k^{e_k}) = (-1)^{e_1 + e_2 + \dots + e_k}$$

where p_1, \dots, p_k are primes.

It is a well known and very deep question whether the set $A = \{n \in \mathbb{N} | \lambda(n) = -1\}$ forms a normal set, see [6] and [7]. It seems that at present we are far away from resolving this outstanding problem. But just for clarity, if the answer for the question is positive, then the aforementioned set A gives us an example of a normal set with no solution to the equation $xy = z$.

For the following we will use a modified Liouville's function λ_Q which is defined by random choice of subset Q inside P (prime numbers) as follows

$$\lambda_Q(p_1^{e_1} p_2^{e_2} \dots p_k^{e_k}) = \lambda_Q(p_1)^{e_1} \lambda_Q(p_2)^{e_2} \dots \lambda_Q(p_k)^{e_k}$$

and

$$\lambda_Q(p) = \begin{cases} -1 & p \in Q \\ 1 & p \notin Q \end{cases}$$

By randomness of Q we mean that a choice of every prime number p is independent of other prime numbers and $Pr(p \in Q) = 0.5$ for any $p \in P$.

One defines $A_Q = \{n \in \mathbb{N} | \lambda_Q(n) = -1\}$. We prove the following

Theorem 5.1.1 *For almost every Q the set A_Q is normal.*

This theorem gives us an infinite family of normal sets such that the multiplicative Schur's equation is not solvable in these sets.

In the section (5.3.1) we prove that the equations $xy = z^2$, $x^2 + y^2 = \text{square}$ and $u^2 - v^2 = \text{square}$ are in DSN .

5.2 A_Q is normal for a.e. Q

We start from an obvious claim about normality of A_Q which is a restatement of lemma 6.3.

Lemma 5.2.1 *Let $Q \subset P$ be given, then A_Q is a normal set \Leftrightarrow for any $k \in (\mathbb{N} \cup \{0\})$ and any $i_1 < i_2 < \dots < i_k$ we have*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \lambda_Q(n) \lambda_Q(n + i_1) \dots \lambda_Q(n + i_k) = 0.$$

Denote

$$T_N = \sum_{n=1}^N \lambda_Q(n) \lambda_Q(n + i_1) \dots \lambda_Q(n + i_k). \quad (5.1)$$

The next step is to show

$$\sum_{N=1}^{\infty} E\left(\left(\frac{1}{N^{40}} \sum_{n=1}^{N^{40}} \lambda_Q(n) \lambda_Q(n + i_1) \dots \lambda_Q(n + i_k)\right)^2\right) < \infty.$$

Lemma 5.2.2 *With T_N as defined in (5.1), $E(T_N^2) \leq O(\frac{1}{N^{0.05}})$.*

Proof. By linearity of expectation we get

$$E(T_N^2) = \frac{1}{N^2} \sum_{x,y=1}^N E(\lambda_Q(x) \lambda_Q(x + i_1) \dots \lambda_Q(x + i_k) \lambda_Q(y) \lambda_Q(y + i_1) \dots \lambda_Q(y + i_k)).$$

Note that for any $m \in \mathbb{N}$, $E(\lambda_Q(m)) = 0$ unless m is a square in which case $E(\lambda_Q(m)) = 1$.

Let us denote by

$$\phi(x) \doteq \lambda_Q(x) \lambda_Q(x + i_1) \dots \lambda_Q(x + i_k)$$

and

$$\xi(x) \doteq x(x + i_1) \dots (x + i_k).$$

By distribution of Q we get

$$E(\phi(x) \phi(y)) = 1 \Leftrightarrow \xi(x) \xi(y) = m^2.$$

Otherwise

$$E(\phi(x)\phi(y)) = 0.$$

Therefore, to obtain an upper bound on $E(T_N^2)$, we give an upper bound on the number of pairs $(x, y) \in [1, N] \times [1, N]$ which satisfy $\xi(x)\xi(y) = \text{square}$.

For a given $x \in [1, N]$ let us assume that $\xi(x) = c_x m^2$, where c_x is a square-free number, say $c_x = p_{j_1} \dots p_{j_l}$ is the prime factorization of c_x . Then we will define $h(x) = l$ (thus $h(x)$ is a number of primes in prime factorization of maximal square-free number which divides x). Denote by D the set of all possible common divisors of the numbers $x, x+i_1, \dots, x+i_k$ (i.e. positive integers which divide at least two of them). For a finite non empty set S of positive numbers we denote by $m(S)$ the product of all elements of S and, for empty set, we fix $m(\emptyset) = 1$.

Note that $\xi(x)\xi(y) = \text{square} \Rightarrow$ there exist $S_1 \subset D$ and $S_2 \subset \{p_{j_1}, \dots, p_{j_l}\}$ such that $y = m(S_1)m(S_2)\text{square}$.

Assume $|D| = r$ (r depends only on the set $\{i_1, \dots, i_k\}$ and doesn't depend on x). Then we obtain $\xi(x)\xi(y) = \text{square}$ for at most $2^r 2^{h(x)} \sqrt{N}$ y 's inside $[1, N]$. Thus

$$E(T_N^2) \leq \frac{1}{N^2} \left(\sum_{n=1}^N 2^r 2^{h(n)} \sqrt{N} \right) \leq \frac{c}{N^{1.5}} \sum_{n=1}^N 2^{h(n)}$$

Therefore it remains to bound the expression $\sum_{n=1}^N 2^{h(n)}$.

If $\xi(n)$ does not contain as dividers 2, 3 then $h(n) \leq \log_5 (n + i_k)^{k+1} = (k+1) \frac{\log_2 (n+i_k)}{\log_2 5}$.

This gives us

$$2^{h(n)} \leq 2^{k+1} (n + i_k)^{\frac{1}{\log_2 5}} \leq C_1 (n + i_k)^{0.45}$$

But if $\xi(n)$ contains 2 or 3 as dividers then $h(n)$ can increase by at most two, this means $2^{h(n)} \leq 2^2 C_1 (n + i_k)^{0.45}$. Thus $\sum_{n=1}^N 2^{h(n)} \leq C_2 (N + i_k)^{1.45}$ and therefore we get

$$E(T_N^2) \leq C_3 \frac{1}{N^{0.05}}.$$

□

Proof. (theorem 5.1.1) From the last lemma we conclude that $\sum_{N=1}^{\infty} E(T_{N^{40}}^2) < \infty$. Thus almost surely $T_{N^{40}} \rightarrow 0$. By lemma 6.4 it follows that almost surely $T_N \rightarrow 0$. And from lemma 5.2.1 (and countability of necessary conditions) it follows that for almost all $Q \subset P$ the sets A_Q are normal.

□

We can now demonstrate the main result of this note.

Theorem 5.2.1 *There exists a normal set $A \subset \mathbb{N}$ such that the multiplicative Schur's equation is not solvable inside A .*

Proof. We have already shown the existence of many Q ($Q \subset P$) such that A_Q are normal. By definition of A_Q follows that for any $x, y \in A_Q$ the number $xy \notin A_Q$. Therefore we can't find $x, y, z \in A_Q$ such that $xy = z$.

□

Corollary 5.2.1 *For any equation $xy = cn^k$ (where c, k are natural numbers, c is not a square and k is even) we can find a normal set $A_{c,k} \subset \mathbb{N}$ such that for any $x, y \in A$ we have $xy \neq cn^k$ for every natural n .*

Proof. We take A_Q be a normal and such that $\lambda_Q(c) = -1$ (it happens with the positive probability $\frac{1}{2}$, and thus there exist such sets). Then obviously we can't solve the proposed equation inside A_Q .

□

5.3 Solvability of the equation $xy = z^2$ and related problems

Theorem 5.3.1 *(theorem 1.4.4 of §1.4.3)*

Let $A \subset \mathbb{N}$ be a WM set. Then there exist $x, y, z \in A$ ($x \neq y$) such that $xy = z^2$.

Proof. For a set $S \subset \mathbb{N}$ let us define $S_a = \{n \in \mathbb{N} | an \in S\}$, where $a \in \mathbb{N}$. It is easily seen that if S is a WM set then S_a is again a WM set with the same statistics as S for any natural a (see [9]). We denote by $d(S)$ density of a set S , if it exists.

Let A be a WM set. We denote by $R_n \doteq A_{2^n}$. For any n , $d(R_n) = \frac{1}{2}$. Let us denote by

$$\mu_N(S) = \frac{|S \cap \{1, 2, \dots, N\}|}{N}$$

for any $S \subset \mathbb{N}$ and any $N \in \mathbb{N}$.

By Szemerédi's theorem (finite version), for any $\delta > 0$ and any $l \in \mathbb{N}$ there exists $N(l, \delta)$ such that for any $N \geq N(l, \delta)$ and any $F \subset \{1, 2, \dots, N\}$ such that $\frac{|F|}{N} \geq \delta$ the set F contains an arithmetic progression of length l (see [16]).

One chooses $K \geq N(3, \frac{1}{3})$. Then there exists N_K such that $\mu_{N_K}(R_i) \geq \frac{1}{3}$ for every $1 \leq i \leq K$.

We claim that there exists $F \subset \{1, 2, \dots, K\}$ such that $\frac{|F|}{K} \geq \frac{1}{3}$ and $\mu_{N_K}(\cap_{j \in F} R_j) > 0$. If not, let us denote 1_{R_i} to be the indicator function of the set R_i inside the set $\{1, \dots, N_K\}$. Then

$$\int_{[1, N_K]} (1_{R_1} + \dots + 1_{R_K}) d\mu_{N_K} = \sum_{j=1}^K \int_{[1, N_K]} 1_{R_j} d\mu_{N_K} \geq \frac{K}{3}.$$

Therefore $\exists n : 1 \leq n \leq N_K$ such that $\sum_{j=1}^K 1_{R_j}(n) \geq \frac{K}{3}$.

Thus $\mu_{N_K}(\cap_{j \in F} R_j) > 0$.

Let $F \subset \{1, 2, \dots, K\}$ such that $\frac{|F|}{K} \geq \frac{1}{3}$ and $\mu_{N_K}(\cap_{j \in F} R_j) > 0$. Then by the choice of K it follows that F necessarily contains arithmetic progression of length 3. The last statement means there exist $a, b, c \in F$ such that $a + c = 2b$. Let us take R_a, R_b, R_c . We have $R_a \cap R_b \cap R_c \neq \emptyset$ and this means there exists $n \in \mathbb{N}$ such that $n2^a \in A$ and $n2^b \in A$ and $n2^c \in A$. Let us denote by x, y, z the following elements of A : $x = n2^a$, $y = n2^c$, $z = n2^b$. Then we have

$$xy = z^2.$$

□

Question: Are the equations $xy = c^2z^2$, where $c > 0$ is a natural number, always solvable inside an arbitrary normal set?

We repeat the formulation of theorem 1.4.5.

Theorem 5.3.2 *Let $A \subset \mathbb{N}$ be an arbitrary normal set. Then there exist $x, y, u, v \in A$ such that $x^2 + y^2 = \text{square}$ and $u^2 - v^2 = \text{square}$.*

Proof. Note that there exist $a, b, c \in \mathbb{N}$ such that $a^2 + b^2 = \text{square}$ and $a^2 + c^2 = \text{square}$ and $b^2 + c^2 = \text{square}$. For example $a = 44, b = 117, c = 240$.

Let $A \subset \mathbb{N}$ be an arbitrary normal set. We look at A_a, A_b, A_c which are defined as in the proof of theorem 5.3.1. Then $d(A_a) = d(A_b) = d(A_c) = \frac{1}{2}$ and thus it can not be true that the intersection of each pair from the triple is empty.

Without loss of generality, let us assume that $A_a \cap A_b \neq \emptyset$.

Thus there exists $z \in A_a \cap A_b$ or equivalently $za, zb \in A$. But $a^2 + b^2 = \text{square}$ and therefore $(za)^2 + (zb)^2 = \text{square}$.

A proof that the equation $u^2 - v^2 = \text{square}$ is solvable in any normal set is similar. We use the fact that there exist $a, b, c \in \mathbb{N}$ with $a < b < c$ such that $c^2 - b^2 = \text{square}$ and $c^2 - a^2 = \text{square}$ and $b^2 - a^2 = \text{square}$. For example $a = 153, b = 185, c = 697$.

□

Question: For an arbitrary normal set A do there exist $x, y, z \in A$ such that $x^2 + y^2 = z^2$?

6 Appendix

In this section we prove all technical lemmas and propositions that were used in the thesis.

We start with the key lemma which is a finite modification of Bergelson's lemma in [2] and its origin is in lemma of van der Corput.

Lemma 6.1 (*van der Corput*) *Suppose $\varepsilon > 0$ and $\{u_j\}_{j=1}^\infty$ is a family of vectors in Hilbert space, such that $\|u_j\| \leq 1$ ($1 \leq j \leq \infty$). Then there exists $I'(\varepsilon) \in \mathbb{N}$, such that for every $I \geq I'(\varepsilon)$ there exists $J'(I, \varepsilon) \in \mathbb{N}$, such that the following holds: For $J \geq J'(I, \varepsilon)$ for which we obtain*

$$\left| \frac{1}{J} \sum_{j=1}^J \langle u_j, u_{j+i} \rangle \right| < \frac{\varepsilon}{2},$$

for set of i 's in the interval $\{1, \dots, I\}$ of density $1 - \frac{\varepsilon}{3}$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^J u_j \right\| < \varepsilon.$$

Proof. For an arbitrary J define $u_k = 0$ for every $k < 1$ or $k > J$. The following is an elementary identity:

$$\sum_{i=1}^I \sum_{j=1}^{J+I} u_{j-i} = I \sum_{j=1}^J u_j.$$

Therefore, the inequality $\left\| \sum_{i=1}^N u_i \right\|^2 \leq N \sum_{i=1}^N \|u_i\|^2$ yields

$$\begin{aligned} \left\| I \sum_{j=1}^J u_j \right\|^2 &\leq (J+I) \sum_{j=1}^{J+I} \left\| \sum_{i=1}^I u_{j-i} \right\|^2 = \\ &(J+I) \sum_{j=1}^{J+I} \langle \sum_{p=1}^I u_{j-p}, \sum_{s=1}^I u_{j-s} \rangle = \\ &(J+I) \sum_{j=1}^{J+I} \sum_{p=1}^I \|u_{j-p}\|^2 + 2(J+I) \sum_{j=1}^{J+I} \sum_{r,s=1; r < s}^I \langle u_{j-r}, u_{j-s} \rangle = \\ &(J+I)(\Sigma_1 + 2\Sigma_2), \end{aligned}$$

where $\Sigma_1 = I \sum_{j=1}^J \|u_j\|^2$ by the aforementioned elementary identity and $\Sigma_2 = \sum_{h=1}^{I-1} (I-h) \sum_{j=1}^J \langle u_j, u_{j+h} \rangle$. The last expression is obtained by rewriting Σ_2 , where $h = r - s$. By dividing the foregoing inequality by $I^2 J^2$ we obtain

$$\left\| \frac{1}{J} \sum_{j=1}^J u_j \right\|^2 < \frac{J+I}{IJ} + \frac{J+I}{J} \left(\frac{\varepsilon}{2} + \frac{\varepsilon}{3} \right) = \frac{J+I}{J} \left(\frac{1}{I} + \frac{5\varepsilon}{6} \right).$$

Choose $I'(\varepsilon) \in \mathbb{N}$, such that $\frac{12}{\varepsilon} \leq I'(\varepsilon) \leq \frac{12}{\varepsilon} + 1$. Then for every $I \geq I'(\varepsilon)$ we have $\frac{1}{I} + \frac{5\varepsilon}{6} \leq \frac{11\varepsilon}{12}$. There exists $J'(I, \varepsilon) \in \mathbb{N}$, such that for every $J \geq J'(I, \varepsilon)$ we obtain $\frac{J+I}{J} < \frac{12}{11}$. As a result, for every $I \geq I'(\varepsilon)$ there exists $J'(I, \varepsilon)$, such that for every $J \geq J'(I, \varepsilon)$

$$\left\| \frac{1}{J} \sum_{j=1}^J u_j \right\|^2 < \varepsilon.$$

□

The next proposition is useful in section 3.1.

Proposition 6.1 *Let $A \subset \mathbb{N}$ be a WM-set. Then for every integer $a > 0$ and every integers b_1, b_2, \dots, b_k we obtain the following*

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \xi(n + b_1) \xi(n + b_2) \dots \xi(n + b_k) = \\ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \xi(an + b_1) \xi(an + b_2) \dots \xi(an + b_k), \end{aligned}$$

where $\xi \doteq 1_A - d(A)$.

Proof. Consider the weak-mixing measure preserving system $(X_\xi, \mathbb{B}, \mu, T)$. The left side of the equation in the proposition is $\int_{X_\xi} T^{b_1} f T^{b_2} f \dots T^{b_k} f d\mu$, where $f(\omega) \doteq \omega_0$ for every infinite sequence inside X_ξ . We make use of the notion of disjointness of measure preserving systems. By [9] we know that every weak-mixing system is disjoint from any Kronecker system which is a compact monothetic group with Borel σ -algebra, the Haar probability measure, and the shift by an a priori chosen element of the group. In particular, every weak-mixing system is disjoint from the measure preserving system $(\mathbb{Z}_a, \mathbb{B}_{\mathbb{Z}_a}, S, \nu)$, where $\mathbb{Z}_a = \mathbb{Z}/a\mathbb{Z}$, $S(n) \doteq n + 1 \pmod{a}$. The measure and the σ -algebra of the last system are uniquely determined. Therefore, from Furstenberg's theorem (see [9]) it follows that the point $(\xi, 0) \in X_\xi \times \mathbb{Z}_a$ is a generic point of the product system $(X_\xi \times \mathbb{Z}_a, \mathbb{B} \times \mathbb{B}_{\mathbb{Z}_a}, T \times S, \mu \times \nu)$. Thus, for every continuous function g on $X_\xi \times \mathbb{Z}_a$ we obtain

$$\int_{X_\xi \times \mathbb{Z}_a} g(x, m) d\mu(x) d\nu(m) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N g(T^n \xi, S^n 0).$$

Let $g(x, m) \doteq f(x) 1_0(m)$ which is obviously continuous on $X_\xi \times \mathbb{Z}_a$. Then genericity of the point $(\xi, 0)$ yields

$$\begin{aligned} \int_{X_\xi \times \mathbb{Z}_a} f(x) 1_0(m) d\mu(x) d\nu(m) &= \frac{1}{a} \int_{X_\xi} f(x) d\mu(x) = \\ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(T^n \xi) 1_0(n) &= \lim_{N \rightarrow \infty} \frac{1}{a} \frac{1}{N} \sum_{n=1}^N f(T^{an} \xi). \end{aligned}$$

Taking instead of the function f the continuous function $T^{b_1}fT^{b_2}f\ldots T^{b_k}f$ in the definition of g finishes the proof.

□

The following lemma is simple fact that for a weak-mixing system X not only an average of shifts for a function converge to a constant in L^2 norm but also weighted averages (weights are bounded) converge to the same constant.

Lemma 6.2 *Let (X, \mathbb{B}, μ, T) be a weak-mixing system and $f \in L^2(X)$ with $\int_X f d\mu = 0$. Let $\varepsilon > 0$. Then there exists $\mathbb{J} > 0$ such that for any $J > \mathbb{J}$ we have*

$$\left\| \frac{1}{J} \sum_{j=1}^J b_j T^j f \right\|_{L^2(X)} < \varepsilon$$

for any sequence $b = (b_1, b_2, \dots, b_n, \dots) \in \{0, 1\}^{\mathbb{N}}$.

Proof. Let $\varepsilon > 0$.

By one of the properties of weak mixing, for any $f \in L^2(X)$ with $\int_X f d\mu(x) = 0$ we have $\frac{1}{N} \sum_{n=1}^N |\langle T^n f, f \rangle| \rightarrow 0$.

We denote by $c_n = c_{(-n)} = |\langle T^n f, f \rangle|$ and we have that $\frac{1}{N} \sum_{n=1}^N c_n \rightarrow 0$. Then for any $\varepsilon > 0$ there exists $\mathbb{J} > 0$ such that for any $J > \mathbb{J}$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^J b_j T^j f \right\|^2 \leq \frac{1}{J^2} \sum_{j=1, k=1}^J b_j b_k c_{j-k} \leq \frac{1}{J^2} \sum_{j=1, k=1}^J c_{j-k} \leq \varepsilon.$$

□

The next two lemmas are very useful for constructing normal sets with specific properties (we use them in this thesis for constructing counterexamples).

Lemma 6.3 *Let $A \subset \mathbb{N}$. Let $\lambda(n) = 21_A(n) - 1$. Then A is a normal set \Leftrightarrow for any $k \in (\mathbb{N} \cup \{0\})$ and any $i_1 < i_2 < \dots < i_k$ we have*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \lambda(n) \lambda(n + i_1) \dots \lambda(n + i_k) = 0.$$

Proof. "⇒" If A is normal then any finite word $w \in \{-1, 1\}^*$ has the "right" frequency $\frac{1}{2^{|w|}}$ inside w_A . This guarantees that "half of the time" the function $\lambda(n) \lambda(n + i_1) \dots \lambda(n + i_k)$ equals 1 and "half of the time" is equal to -1. Therefore we get the desired conclusion.

"⇐" Let w be an arbitrary finite word of plus and minus ones: $w = a_1 a_2 \dots a_k$ and we have to prove that w occurs in w_A with the frequency 2^{-k} . For every $n \in \mathbb{N}$ the word w occurs in 1_A and starting from n if and only if

$$\begin{cases} 1_A(n) = a_1 \\ \dots \\ 1_A(n + k - 1) = a_k \end{cases}$$

The latter is equivalent to the following

$$\begin{cases} \lambda(n) = 2a_1 - 1 \\ \dots \\ \lambda(n + k - 1) = 2a_k - 1 \end{cases}$$

The frequency of w within 1_A is equal to

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \frac{\lambda(n)(2a_1 - 1) + 1}{2} \dots \frac{\lambda(n + k - 1)(2a_k - 1) + 1}{2}.$$

By assumptions of the lemma the latter expression is equal to $\frac{1}{2^k}$.

□

Lemma 6.4 *Let $\{a_n\}$ be a bounded sequence. Denote by $T_N = \frac{1}{N} \sum_{n=1}^N a_n$. Then T_N converges to a limit $t \Leftrightarrow$ there exists a sequence of increasing indices $\{N_i\}$ such that $\frac{N_i}{N_{i+1}} \rightarrow 1$ and $T_{N_i} \rightarrow_{i \rightarrow \infty} t$.*

References

- [1] Bergelson, V. Sets of recurrence of Z^m -actions and properties of sets of differences in Z^m . J. London Math. Soc. (2) 31 (1985), no. 2, 295–304.
- [2] Bergelson, V. Weakly mixing PET. Ergodic Theory Dynam. Systems 7 (1987), no. 3, 337–349.
- [3] Bergelson, V.; Furstenberg, H.; McCutcheon, R. IP-sets and polynomial recurrence. Ergodic Theory Dynam. Systems 16 (1996), no. 5, 963–974.
- [4] Bergelson, V.; Hindman, N. Additive and multiplicative Ramsey theorems in N —some elementary results. Combin. Probab. Comput. 2 (1993), no. 3, 221–241.
- [5] Bergelson, V.; Leibman, A. Polynomial extensions of van der Waerden’s and Szemerédi’s theorems. J. Amer. Math. Soc. 9 (1996), no. 3, 725–753.
- [6] Cassaigne, J.; Ferenczi, S.; Mauduit, C.; Rivat, J.; Sarközy, A. On finite pseudorandom binary sequences. III. The Liouville function. I. Acta Arith. 87 (1999), no. 4, 367–390.
- [7] Cassaigne, J.; Ferenczi, S.; Mauduit, C.; Rivat, J.; Sarközy, A. On finite pseudorandom binary sequences. IV. The Liouville function. II. Acta Arith. 95 (2000), no. 4, 343–359.
- [8] Fish, A. Random Liouville functions and normal sets. Acta Arith. 120 (2005), no. 2, 191–196.
- [9] Furstenberg, H. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. Math. Systems Theory 1 (1967), 1–49.
- [10] Furstenberg, H. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. J. d’ Analys Math. 31 (1977), 204–256.
- [11] Furstenberg, H. Recurrence in Ergodic Theory and Combinatorial Number Theory. Princeton Univ. Press 1981.
- [12] Host, B.; Kra, B. Nonconventional ergodic averages and nilmanifolds. Ann. of Math. (2) 161 (2005), no. 1, 397–488.
- [13] Host, B.; Kra, B. Convergence of polynomial ergodic averages. Israel J. Math. 149 (2005), 1–19.
- [14] Rado, R. Note on combinatorial analysis. Proc. London Math. Soc. 48 (1943), 122–160.
- [15] Schur, I. Über die Kongruenz $x^m + y^m \equiv z^m \pmod{p}$. Jahresbericht der Deutschen Math.-Ver. 25 (1916), 114–117.

- [16] Szemerédi, E. On sets of integers containing no k elements in arithmetic progression. Collection of articles in memory of Juriĭ Vladimirovič Linnik. Acta Arith. 27 (1975), 199–245.